Lecture Note
Health problems, determinants, and trends

(Course No.5300501: 3 credits (3-0-9))

For Master of Public Health (MPH)

College of Public Health Sciences, Chulalongkorn University

By

Dr. Nutta Taneepanicskul

Assistant Professor, College of Public Health Sciences

Chulalongkorn University

11th floor (Room no.1105) Institute Building 3, Soi Chulalongkorn 62,
Phyathai Rd, Pathumwan Bangkok 10330, Thailand

E-mail: nutta.t@chula.ac.th, nutta.taneepanichskul@gmail.com

Introduction

Epidemiology as defined by Last is *"the study of the distribution and determinants of health-related states or events in specified populations and the application of this study to the prevention and control of health problems"*. Traditional epidemiological studies include quantitative and qualitative study designs. Quantitative study designs include observational and interventional methodology.

In this course, student will introduce to concepts of burden of disease, the epidemiologic, demographic, and nutrition transitions, and types and interplay among major determinants (i.e. demographic, physical environmental, behavioral, psycho-socio-economic, and cultural determinants.) and basic concepts in epidemiology related to assessing burden of disease and major public health-related determinants. Both experimental and non-experimental research design will introduce to students for increasing their understanding in health determinant and epidemiological concept. Two-way communication and active learning procedure will apply throughout the course for critical thinking, analytic skills, and application to research practice. Application of epidemiology will be taught trough environmental and occupation epidemiology concept and prevention for non-communicable diseases and communicable disease

"*Concepts and course structure would be acknowledged Dr.Robert S. Chapman who is an epidemiologist and lecturer by his heart and soul."*

Nutta Taneepanichskul, PhD

Assistant Professor
College of Public Health Sciences, Chulalongkorn University

Contents

Course Syllabus (LWP & INHOUSE)..... 1
Behavioral objectives..... 2
Session 1 Introduction and Overview of Health Determinants..... 3
Session 3 Epidemiology: General concepts 11
Session 4 Epidemiological study: Cross sectional study 31
Session 5 Epidemiological study: Case Control study..... 44
Session 7 Epidemiological study: Cohort Study 55
Session 8 Individual evaluation report (10%)..... 70
Session 9 Mid-term examination (Report – 15%) 71
Session 10 Critical paper evaluation / Group presentation 73
Session 11 Environmental and Occupational Epidemiology 75
Session 15 Field Study 80
Power Point Handout..... 81

Course Syllabus (LWP & INHOUSE)

SESSION/ DATE	TOPICS	INSTRUCTOR
SESSION 1	Introduction and Overview Health Determinants	Dr Nutta
SESSION 2	Disease burden: global, regional and national	Prof.Surasak
SESSION 3	Epidemiology: general concepts Disease investigation and surveillance	Dr.Nutta
SESSION 4	Epidemiological study: Cross sectional study	Dr Nutta
SESSION 5	Epidemiological study: Case Control study	Dr Nutta
SESSION 6	Intervention Study: Quasi Experimental study	Prof.Surasak
SESSION 7	Epidemiological study: Cohort Study	Dr Nutta
SESSION 8	Individual evaluation report (10%) International annual CPHS conference	Dr Nutta
SESSION 9	Mid-term examination: Report (15%)	Dr Nutta
SESSION 10	Critical paper evaluation group presentation (15%)	Dr Nutta
SESSION 11	Environmental and Occupation Epidemiology	Dr Nutta
SESSION 12	Epidemiology Health Policy and Planning	Prof. Sathirakorn
SESSION 13	Epidemiology and Prevention for Non Communicable Diseases and communicable disease	Prof.Wiroj (invited speaker)/ Dr.Nutta
SESSION 14	Final Examination (50%)	Dr. Nutta
SESSION 15	Field Study (Tentative) ** To be changed ; based on registered semester*	Dr. Nutta

Behavioral objectives

1. Have working knowledge of burden of disease on a global and regional basis.

2. Recognize the multifactorial etiology of disease in populations.

3. Have ability to critically evaluate the degree of causation inherent in epidemiologic and statistical associations

4. Be familiar with basic epidemiologic study designs, and direct and indirect adjustment of health-related data

5. Can design research based on epidemiology research design

5300501 – Lecture # Session 1 Introduction and Overview of Health Determinants

Learning Objectives

- Understand epidemiological concept
- Describe descriptive and analytical epidemiology
- Understand determinants of health
- List the three components of the epidemiologic triad with concepts of disease occurrence

What's Health?

"[Health is] the state of complete physical, mental and social well-being and not merely the absence of disease or infirmity" (WHO, 1948)

Figure 1 *Definition of Health*

What does WELL-BEING mean?

Regarding to the health definition, health means a person with physical well-being, mental well-being and social well-being. Having said that "well-being", we may need a measurement that could quantify. Therefore, studying health determinant and health indicators will tell how the well-being health mean.

What's Public Health?

Center of disease control (CDC) define "Public health is the science of protecting and improving the health of people and their communities."

American Public Health Association (APHA) define "Public health promotes and protects the health of people and the communities where they live, learn, work and play".

Protecting and improving the health of people can be done if we know the baseline *information of health and the influencing health factors* among population or community. At this point, health indicators play its importance role.

Health Indicators

A health indicator is a variable – that can be measured directly to reflect the state of health of people within a community.

"Health indicators" are a set of surveillance data that has been analysed to show health status of the population. Public health priorities and actions can be appropriately identified after analyzing data from health indicators. Health indicators should be comprehensive, valid (sensitive), standardised, meet quality criteria, and be flexible (never fixed and final) to support evolving health policies.

Regarding to European Core Health Indicators (ECHI)[1], health indicators are divided into

1. Health status indicators (disease specific mortality, morbidity and disability)
2. Determinants of health (individual characteristic / health behaviours and social health determinant)

1. Health status indicators

Disease-specific mortality, morbidity and disability should be available for the overall population (national level and local level) with age and sex specific. Sometimes, data present in absolute numbers, crude and age-standardised rates. Standardization would be benefit for comparability between countries with different age distribution in the population.

1.1 Mortality data provide a crude but simple way to assess health conditions

- Source of information are death certificates where the underlying cause of death is coded.
- The reliability of mortality data depends on the completeness and accuracy of the registration system of the country.

Coding of death is according to the International Classification of Diseases and Causes of Death (ICD). Problems of temporal and geographic comparisons derive from the different versions of the ICD (8th, 9th, 10th and 11th (2018 version)) and from different coding practices in each country

An example of mortality indicators:

i) Crude death rate: the number of overall and cause-specific deaths divided by estimated midyear population per 1,000
ii) Age-standardised death rate: death rate estimated after age-standardisation has been performed
iii) Age-specific death rate: the number of deaths divided by estimated mid-year population per 1,000 for specific age groups

1.2 Morbidity data describe using the following frequency measurements:

- Hospital discharge rate
- Attack/incidence rate
- Prevalence
- Case fatality

[1] Kramers, P. G. (2003). The ECHI project: health indicators for the European Community. *The European Journal of Public Health*, *13*(supplement 1), 101-106.

An example of morbidity indicators:

i) Hospital discharge rate or hospitalisation rate: the number of hospital discharge records of a specified main diagnosis divided by the estimated mid-year population per 1,000
ii) Attack rate: the number of events (first and recurrent) divided by the estimated mid-year population per 1,000
iii) Incidence rate: the number of first events divided by the population at risk per 1,000 or person/years at risk. *(More explanation in session 3)*

> Person/years at risk consists of the sum of periods of time (years) at risk contributed by each of the person included in the study. Incidence may be estimated through the follow-up of a population enrolled in a cohort or the identification of new events in a dynamic population. It can be obtained using longitudinal studies or disease registers, when it is possible to eliminate those who have already experienced an event

iv) Prevalence: the proportion of persons with the disease in a population per 100 or per 1000 at a particular time. *(More explanation in session 3)*
v) Case fatality: the number of fatal cases divided by the total number of events. It is usually expressed as the percentage of persons diagnosed as having a specified disease who die from that disease within a given period. *(More explanation in session 3)*

Note: Choosing morbidity indicators may focus on characteristic of interested population. For example, incidence, case fatality and prevalence are important in younger age group. In elderly, attack rate and prevalence are more important since patients with chronic diseases. Incidence rates measure acute event while prevalence are more focus on chronic events.

1.3 Disability data describe using the following frequency measurements.

An example of morbidity indicators:

i.) Disability-Adjusted Life Year (DALY): takes into account years lost due to premature mortality and years lived with disability. It is equal to the sum of the number of years of life lost due to specific diseases in a population (YLL) and the number of years lived with disability of known severity and duration for disease in a population (YLD).
ii.) Potential Years of Life Lost (PYLL): it is a measure of the impact of premature mortality on the population. PYLL is the sum of the years that

people dying from a certain disease would have lived, had they experienced a normal life expectancy.
(More explanation in session 3)

2. Determinants of health

Many factors combine together to affect the health of individuals and communities. Whether people are healthy or not, is determined by their circumstances and environment. To a large extent, factors such as where we live, the state of our environment, genetics, our income and education level, and our relationships with friends and family all have considerable impacts on health, whereas the more commonly considered factors such as access and use of health care services often have less of an impact[1].

Determinants of health are also lifestyle-based properties affected by broader social, economic, and political forces that influence quality of personal health. These attributes include but are not limited to education level, employment, income level and distribution, housing, childhood development, food security and nutrition, race, gender, and stress. Such factors have been shown to have marked associations with risks for different illnesses, life expectancy, and life-time morbidity[2]

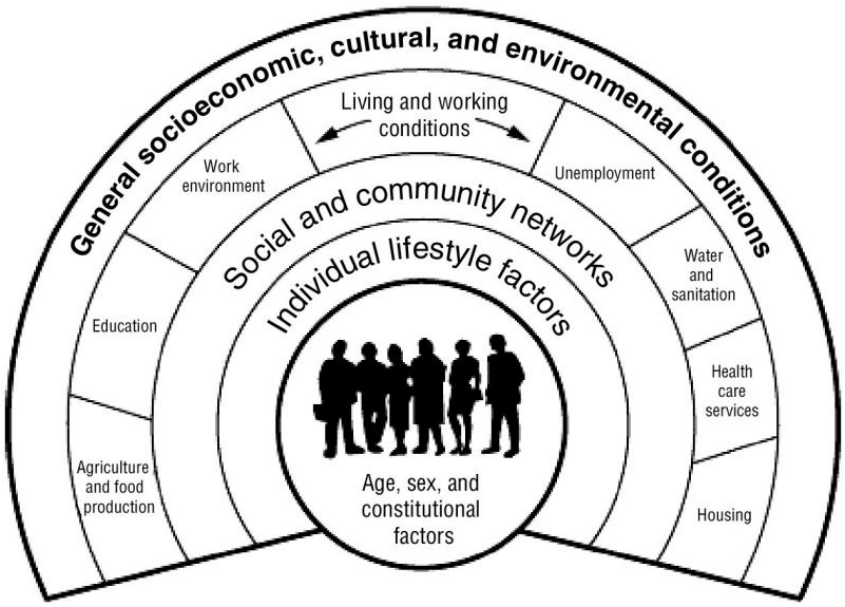


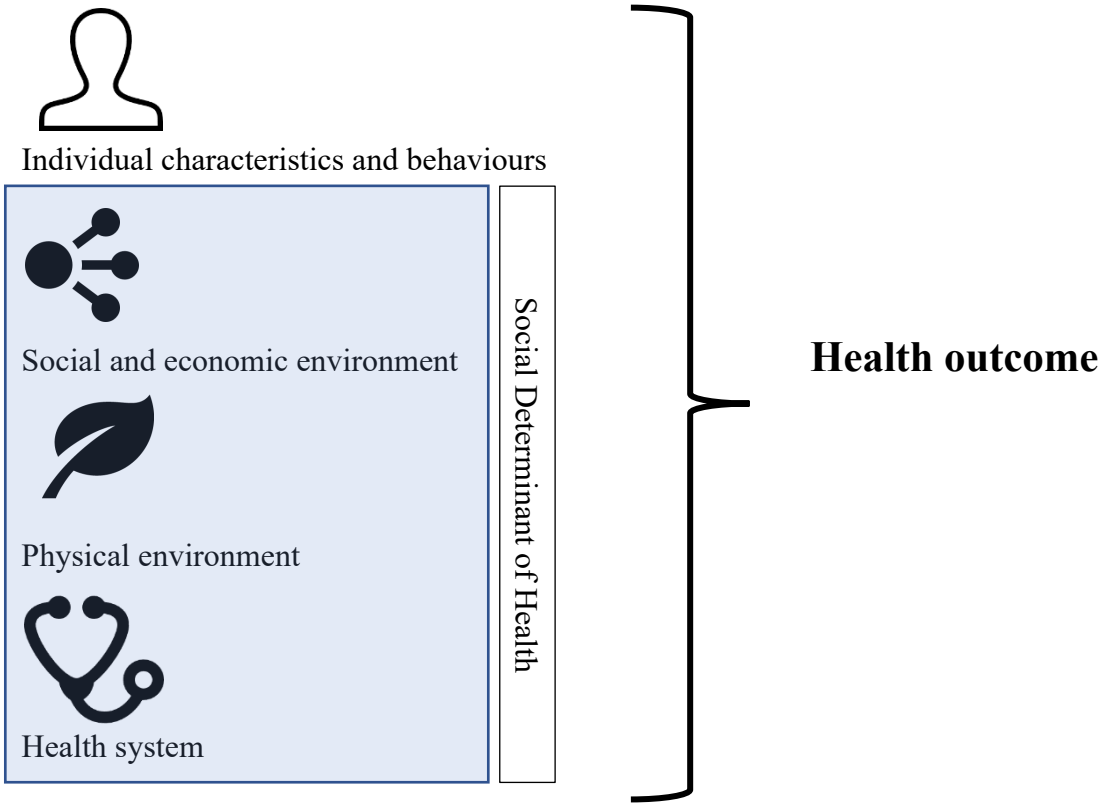
Figure 2 *The Dahlgren and Whitehead model of health determinants*[3]

[1] World Health Organization definition in Health Impact Assessment
[2] Braveman PA, Cubbin C, Egerter S, Williams DR. 2011. The social determinants of health: coming of age. Annu. Rev. Public Health 32:3.1-3.18
[3] Dahlgren G, Whitehead M. 1991. Policies and Strategies to Promote Social Equity in Health. Stockholm, Sweden: Institute for Futures Studies.

From figure 2 indicates that[1]:

- Personal characteristics occupy the core of the model and include sex, age, ethnic group, and hereditary factors.
- Individual 'lifestyle' factors include behaviours such as smoking, alcohol use, and physical activity
- Social and community networks include family and wider social circles
- Living and working conditions include access and opportunities in relation to jobs, housing, education and welfare services
- General socioeconomic, cultural and environmental conditions include factors such as disposable income, taxation and availability of work

To summarize basic idea of health determinants, it can be divided into 4 main aspects as following;



[1] Public Health England, Chapter 6: social determinants of health (July, 2017) Available online: https://www.gov.uk/government/publications/health-profile-for-england/chapter-6-social-determinants-of-health.

Social Determinant of Health

The social determinants of health are the conditions in which people are born, grow, live, work and age. These circumstances are shaped by the distribution of money, power and resources at global, national and local levels. The social determinants of health are mostly responsible for health inequities.

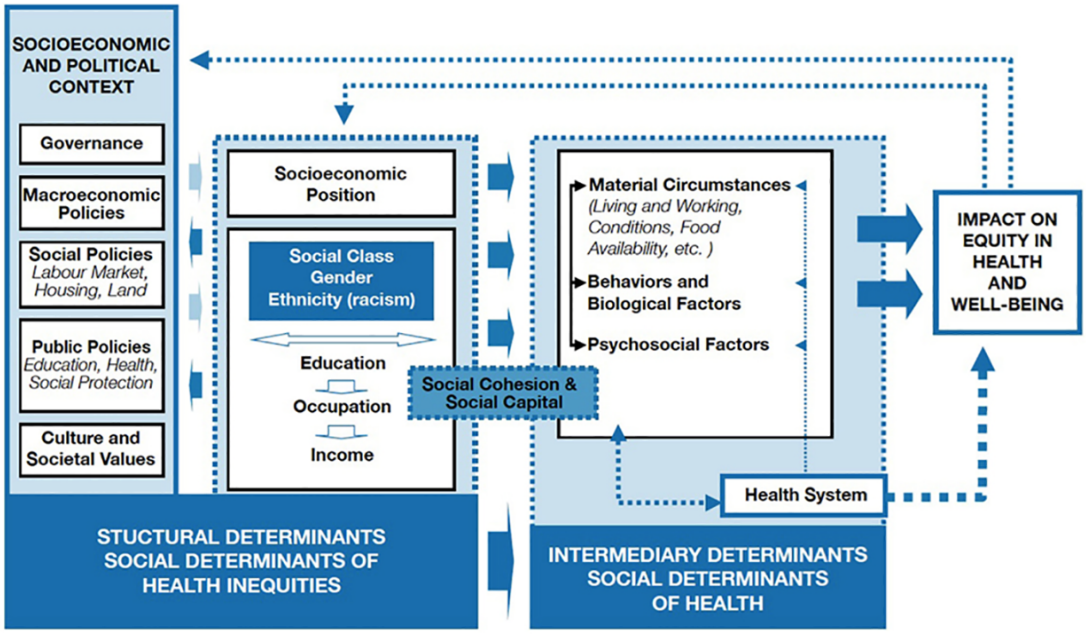


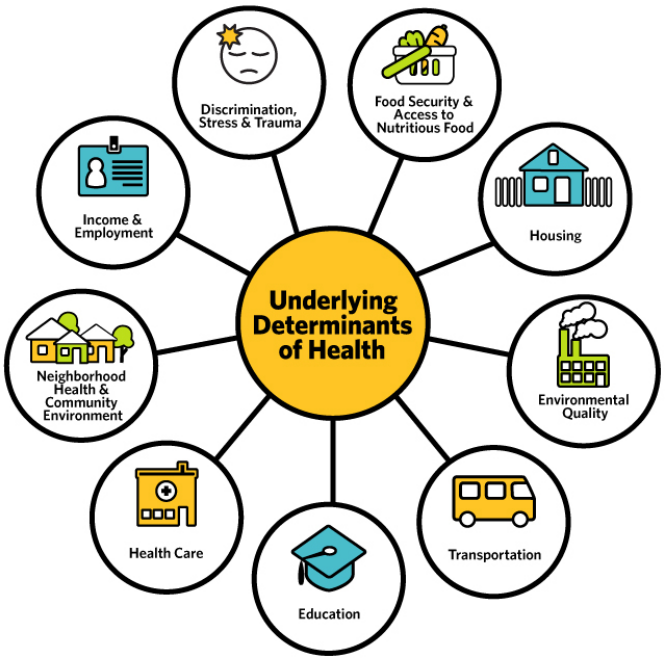
Figure 3 *WHO Social Determinants of Health Framework*[1]

Examples of social determinants:

- Availability of resources to meet daily needs (e.g., safe housing and local food markets)
- Access to educational, economic, and job opportunities
- Access to health care services
- Quality of education and job training
- Availability of community-based resources in support of community living and opportunities for recreational and leisure-time activities
- Transportation options
- Public safety
- Social support

[1] Schröders J, Wall S, Kusnanto H, Ng N (2015) Millennium Development Goal Four and Child Health Inequities in Indonesia: A Systematic Review of the Literature. PLoS ONE 10(5): e0123629. doi:10.1371/journal.pone.0123629

- Social norms and attitudes (e.g., discrimination, racism, and distrust of government)
- Exposure to crime, violence, and social disorder (e.g., presence of trash and lack of cooperation in a community)
- Socioeconomic conditions (e.g., concentrated poverty and the stressful conditions that accompany it)
- Residential segregation
- Language/Literacy
- Access to mass media and emerging technologies (e.g., cell phones, the Internet, and social media)
- Culture



5300501 – Lecture # Session 3 Epidemiology: General concepts

(Measuring Disease and Health)

Learning Objectives

- Explain the differences among the parameters: ratio, proportion, & rate
- Calculate prevalence and incident
- Distinguish between cumulative incidence and incidence rate
- Explain what is meant by "person-years" of observation and be able to calculate person-years

What does "EPIDEMIOLOGY" mean?

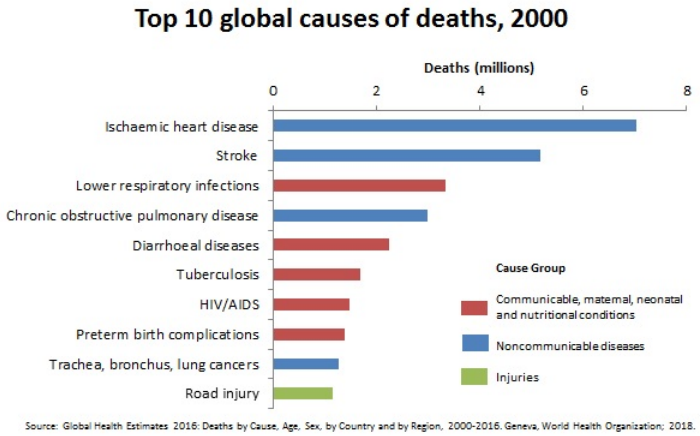
Epi : upon, among

Demos : people

Ology : science, study of

Epidemiology : the science or the study of epidemic

"Epidemiology means the study of distribution and determinants of health, disease, or injury in human populations and the application of this study to the control of health problems."



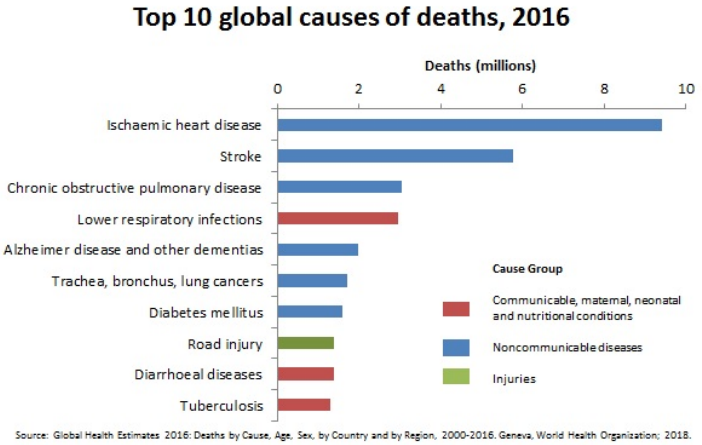


Figure 4 *Comparison of global cause of death in 2000 and 2016 (An example of Descriptive epidemiology)*[1]

[1] World Health Organization, 2018 Available online : http://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

Table 1 Epidemiology definition and epidemiology type[1]

Term	Explanation	Type of Epidemiology
Study	Surveillance, observation, hypothesis testing, analytic research and experiments.	
Distribution	An analysis of times, persons, places and classes of people affected.	Descriptive Epidemiology - Frequency of health events - By person, time and place
Determinants	Include factors that influence health (biological, chemical, physical, social, cultural, economic, genetic and behavioural)	Analytical Epidemiology - Search for causes or risk factors - Response to a study hypothesis - Use various epidemiologic methods
Health-related states and events	Include diseases, causes of death, behaviours such as use of tobacco, positive health states, reactions to preventive regimes and provision and use of health services.	Disease-specific Epidemiology − All health outcomes
Specified populations	Include identifiable characteristics, such as occupational groups	
Application to prevention and control	Aims of public health for promoting, protecting, and restore health.	Applied Epidemiology

[1] Adapted from Bonita, R., Beaglehole, R., & Kjellström, T. (2006). *Basic epidemiology*. World Health Organization and Kanchanaraksa, S. Epidemiologic Investigation, Johns Hopkins University (Openware)

Epidemiologic Triangle (Triad) and Disease Causation

Epidemiologic Triangle is a traditional model for infectious disease before applying to non-infectious disease (figure 5). It composed of an external agent, a susceptible host, and an environment that brings the host and agent together. Agent, host, and environmental factors interrelate in a variety of complex ways to produce disease. Different diseases require different balances and interactions of these three components. Therefore, an appropriate, practical, and effective public health measures to control or prevent disease usually requires assessment of all three components and their interactions.

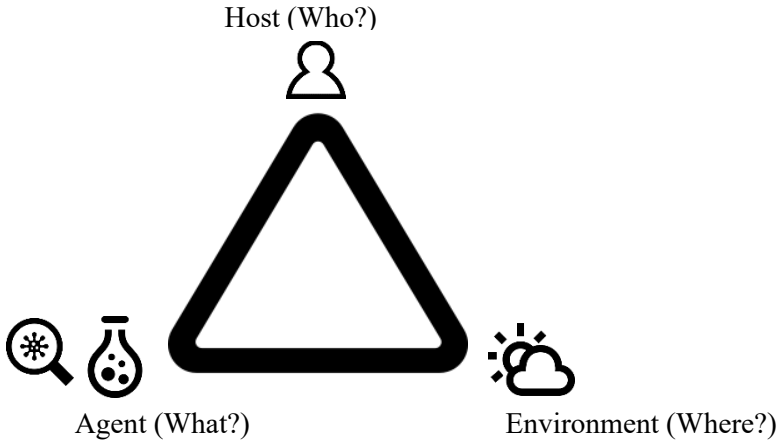


Figure 5 *Epidemiology Triad (developed by Nutta Taneepanichskul)*

Agent (What?)

Agent is originally refer to a microorganism or pathogen which could link to infectious disease e.g. virus, bacteria or parasite. Presently, agent concept is included chemical and physical causes of disease or injury that can be harm to human health. Present of only agent itself could not link to cause disease. Dose and exposure also need to be concern for development of disease. However, epidemiology triad is not adequate for proving some disease such as cardiovascular disease, cancer, and other diseases that appear to have multiple contributing causes without a single necessary one.

Factors related to agent:

- Biological (bacteria, etc.)
- Chemical (poison, alcohol, smoke)
- Physical (auto, radiation, fire)
- Nutritional (lack or excess)

Host (Who?)

Host refers to the human who can get the disease. A variety of factors essential for host, sometimes called risk factors, can influence an individual's exposure, susceptibility, or response to a causative agent. Opportunities for exposure are influenced by behaviors such as hygiene practice and other personal choices as well as by genetic, age, and sex.

Factors related to host:

- Age
- Gender
- Ethnicity
- Religion
- Customs
- Occupation
- Heredity
- Marital status
- Family background
- Previous diseases

Environment (Where?)

Environment refers to factors that affect the agent and the opportunity for exposure. Environmental factors include physical factors such as geology and climate, biologic factors such as socioeconomic factors such as crowding, sanitation, and the availability of health services.

Factors related to environment:

- Temperature
- Humidity
- Altitude
- Crowding
- Housing
- Neighborhood
- Water
- Milk
- Food
- Radiation
- Air pollution
- Noise

Epidemiology and Public Health[1]

Epidemiology is considered the key scientific of public health practice. This role of epidemiology was aimed at the prevention of disease and promotion of health, with linkages to many disciplines and a scientific core of epidemiology.

Epidemiology is the study of the health of human populations. Its functions are as follows:

- *To discover the agent, host, and environmental factors that affect health* to provide the scientific basis for the prevention of disease and injury and the promotion of health;
- *To determine the relative importance of causes of illness, disability, and death* to establish priorities for research and action;
- *To identify those sections of the population that have the greatest risk from specific causes of ill health* so that the indicated action may be directed appropriately
- *To evaluate the effectiveness of health programs and services in improving the health* of the population.

From the concept of epidemiology in public health, natural of disease and causation of illness should be clarify.

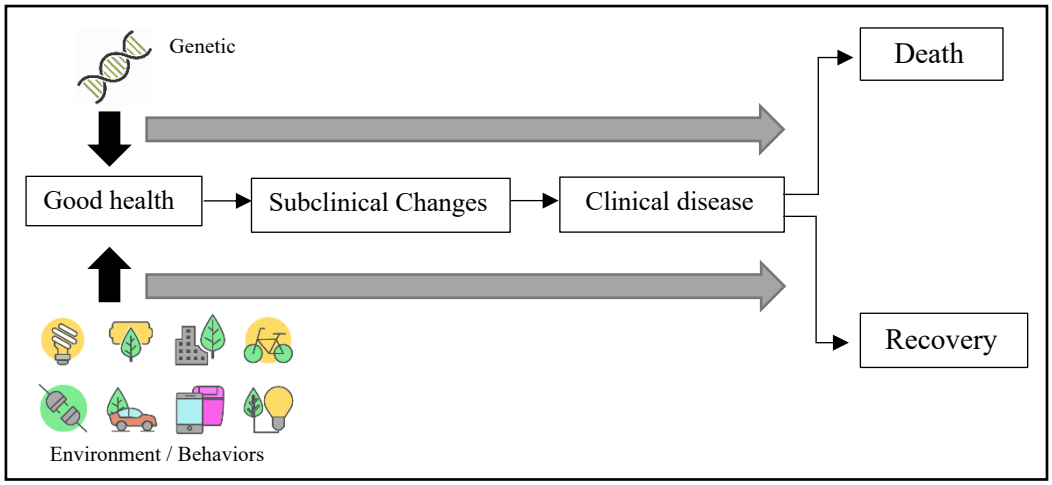


Figure 6 *Natural of disease development and causation of disease*

Genetic, environmental conditions and human behaviors can drive human good health to become ill-health. After influencing by those factors, good health will present some clinical change and lead to clinical disease. Either good health or subclinical changes, genetic, environment and behaviors can drive human health to clinical disease.

[1] Haveman-Nies, A., Jansen, S. C., van Oers, J. A. M., & van't Veer, P. (2011). Epidemiology in Public Health Practice. Am J Epidemiol, 174(7), 871-872.

Regarding to functions of epidemiology in public health, some specific definition should be defined.

Population at risk

The people who are susceptible to a given disease are called the population at risk, and can be defined by demographic, geographic or environmental factors.

- Let's give an easy example. Who are the population at risk of cervical cancer?

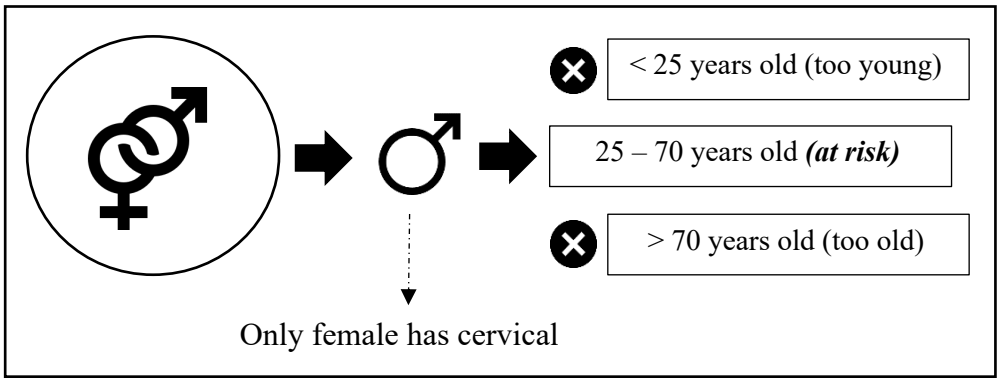


Figure 7 *An example of population at risk*[1]

Endemic, Epidemic, and Pandemic

Endemic − The habitual presence of a disease within a given geographic area − May also refer to the usual prevalence of a given disease within such an area.

Epidemic − The occurrence in a community or region of a group of illnesses of similar nature, clearly in excess of normal expectancy − Outbreak

Pandemic − A worldwide epidemic

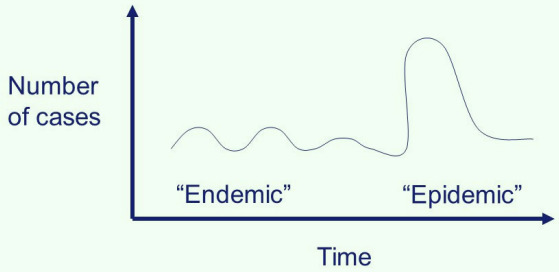
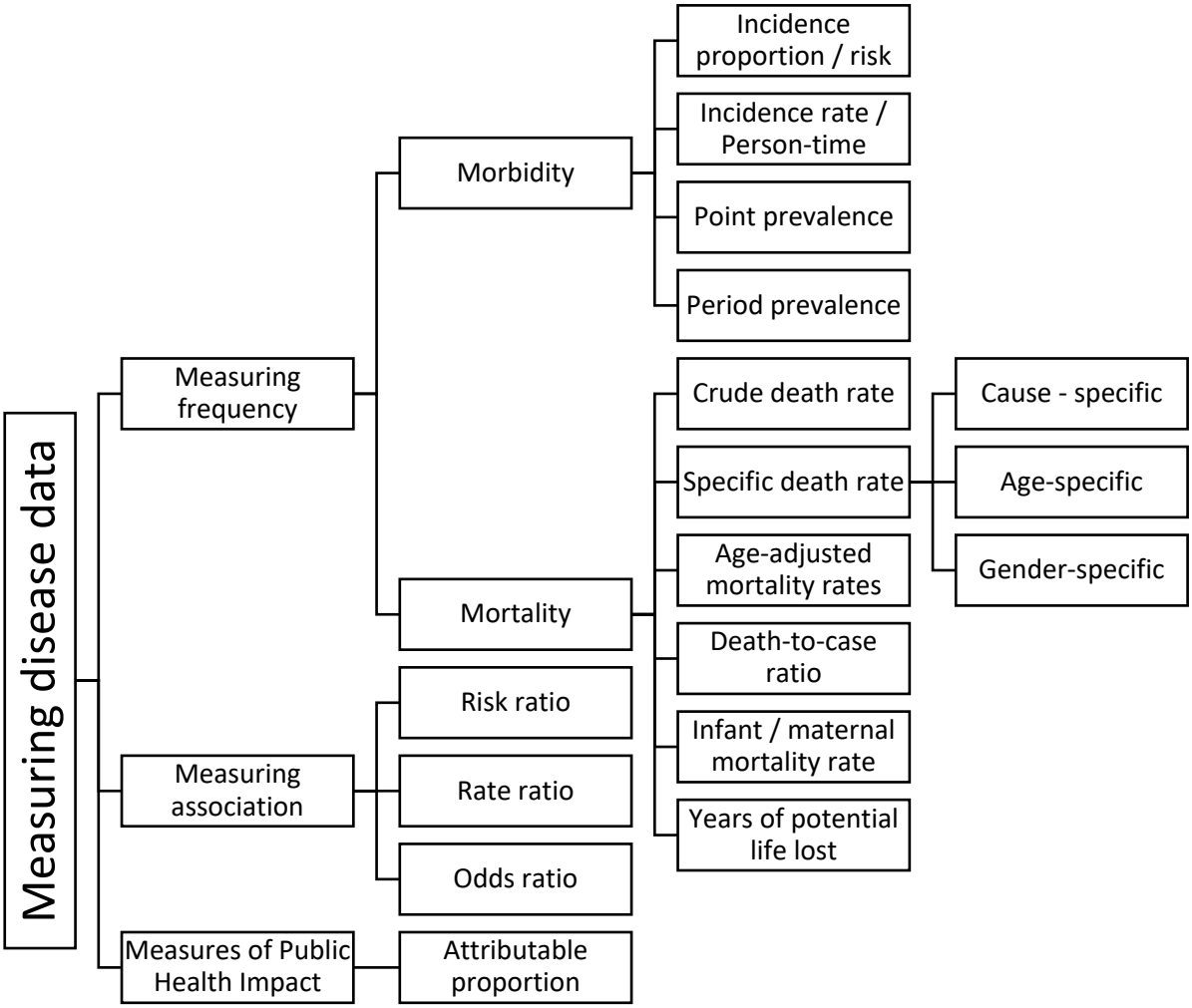


Figure 8 *Endemic and Epidemic of Disease*

[1] Modified from Bonita, R., Beaglehole, R., & Kjellström, T. (2006). *Basic epidemiology*. World Health Organization

Measuring occurrence of disease



Developed by Nutta Taneepanichskul

Measuring occurrence of disease[1]

Common frequency measures are ratios, proportions, and rates. All three frequency measures have the same basic form;

$$\frac{\text{Numerator}}{\text{Denominator}} \times 10^n$$

Ratio is one number divided by another number (numerator doesn't have to be included in denominator—and vice versa)

$$\textbf{Ratio} = \frac{\text{Number or rate of events, items, persons, etc. in one group}}{\text{Number or rate of events, items, persons, etc. in another group}}$$

Example
Ratios can describe the male-to-female ratio of participants in a study, or the ratio of controls to cases.
As an analytic tool, ratios can be calculated for occurrence of illness, injury, or death between two groups.

Proportion is a part of ratio in which the numerator is included in the denominator. It has no dimension because the unit of the numerator cancels out the unit of the denominator

$$\textbf{Proportion} = \frac{\text{Number of persons or events with a particular characteristic}}{\text{Total number of persons or events, of which the numerator is a subset}} \times 10n$$

For a proportion, 10n is usually 100 (or n = 2) and is often expressed as a percentage.

Example

Public health officials have estimated that greater than 90% of the lung cancer cases that occur are attributable to cigarette smoking.

NOTE: the numerator must be included in the denominator.

[1] Centers for Disease Control and Prevention (CDC). An Introduction to Applied Epidemiology and Biostatistics: Principles of Epidemiology in Public Health Practice, (May 2012).

Convert Proportion to Ratio
To convert proportion of male and female who attend clinic
Let's start with number of women who attend clinic is 179.
Total clinic attendees is 341 (male and female)
The proportion of the proportion of clinic attendees who are women is
179 / 341 (52%)
Convert to female-to-male ratio
From total attendees, the number of men is 162 (341 – 179).
Ratio of women to men = 179 / 162 = 1.1 to 1 female-to-male ratio

Rate is a ratio which is a measure of the frequency with which an event occurs in a defined population over a specified period of time. The numerator is the number of events. The unit is event (i.e., number of cases). The denominator is the sum of follow-up time contributed by the people at risk of the event. The unit is time or, more accurately, person-time to account for duration of time of follow-up

Rate = the number of events (i.e., number of cases)
the sum of follow-up time contributed by the people at risk of the event (person-time)

Example

70 new cases of breast cancer per 1,000 women per year

In non-epidemiologist way

Rate means how fast something is happening or going eg. car's speed or rate of travel in miles or kilometers per hour. (reported per some unit of time)

Compare to epidemiologist way

70 new cases of breast cancer per 1,000 women per year (reported per some unit of time) This measure conveys a sense of the speed with which disease occurs in a population. It can imply that this pattern has occurred and will continue to occur for the future.

Measuring Morbidity frequency

Incidence and Prevalence

- Incidence of disease represents the rate of occurrence of new cases arising in a given period in a specified population.
- Prevalence is the frequency of existing cases in a defined population at a given point in time.

Table 2 *Comparison of incidence and prevalence*[1]

	Incidence	Prevalence
Numerator	Number of new cases of disease during a specified period of time Key word : NEW CASE	Number of existing cases of disease at a given point of time Key word : EXISTING CASES
Denominator	Population at risk	Population at risk
Focus	Whether the event is a new case Time of onset of the disease	Presence or absence of a disease Time period is arbitrary; rather a "snapshot" in time

Table 3 *Morbidity measurement*

Table 3.3 Frequently Used Measures of Morbidity

Measure	Numerator	Denominator
Incidence proportion (or attack rate or risk)	Number of new cases of disease during specified time interval	Population at start of time interval
Secondary attack rate	Number of new cases among contacts	Total number of contacts
Incidence rate (or person-time rate)	Number of new cases of disease during specified time interval	Summed person-years of observation or average population during time interval
Point prevalence	Number of current cases (new and preexisting) at a specified point in time	Population at the same specified point in time
Period prevalence	Number of current cases (new and preexisting) over a specified period of time	Average or mid-interval population

[1] Bonita, R., Beaglehole, R., & Kjellström, T. (2006). *Basic epidemiology*. World Health Organization

Incidence proportion or risk

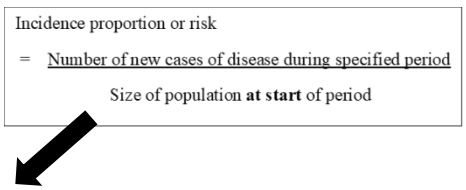
Incidence proportion or risk
= Number of new cases of disease during specified period / Size of population **at start** of period

Incidence proportion is the proportion of an initially disease-free population that develops disease (die) **at specified period of time.**

PROPORTION - the numerator is included in the denominator

Incidence proportion is a proportion because the persons who develop disease are all included in the entire population.

Synonyms for incidence proportion
• Attack rate
• Risk
• Probability of developing disease
• Cumulative incidence


Incidence proportion or risk
= Number of new cases of disease during specified period
Size of population at start of period

Population at start of period (Denominator)

- Should be limited to the "population at risk" for developing disease

For example : the new cases is ovaries cancer patients, the denominator should be restricted to women.

During disease outbreak, the term "**attack rate"** is often used as a synonym for risk. It is the risk of getting the disease during a specified period, such as the duration of an outbreak.

- Overall attack rate is the total number of new cases divided by the total population.
- A food-specific attack rate is the number of persons who ate a specified food and became ill divided by the total number of persons who ate that food

Incidence rate (person-time)

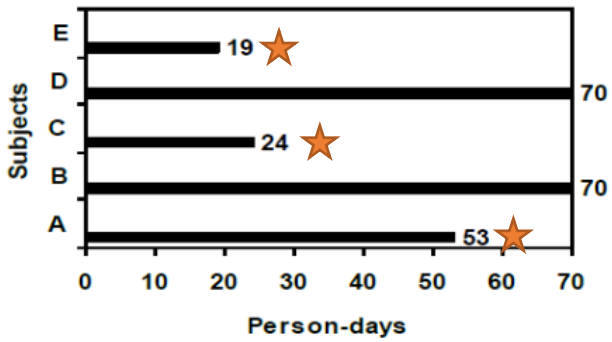
Incidence rate = Number of new cases of disease or injury during specified period Time each person was observed (totaled for all persons)

RATE: The denominator is the sum of follow-up time contributed by the people at risk of the event.

Person-time

- A person-time rate is generally calculated from a long-term cohort follow-up study.
- Person-time is an estimate of the actual time-at-risk (in years (normally use), months, or days) – that all participants contributed to a study.

Example [1]



★ Developed second MI

Suppose one researcher is conducting a study of the rate of second myocardial infarction (MI). She follow 5 patients which have first MI up to 10 weeks. The results is showed above.

X-axis presents number of days for each person contribute to the study.

Y-axis presents 5 patients who were follow up in this study.

From the graph, we can calculate person-days (unit is days) for 5 patients who contribute to the study.

[1] Alexander, L. K., Lopes, B., Ricchetti-Masterson, K., & Yeatts, K. B. Calculating Person-Time.

Time contributed by each subject:

Subject A: 53 days

Subject B: 70 days

Subject C: 24 days

Subject D: 70 days

Subject E: 19 days

Total person-days in the study: 53+70+24+70+19=236 person-days

All ★ means patients who developed second MI. So, there are 3 patients who have recurrent of MI during 10 week (70 days) of follow-up.

Then, we convert to incidence rate.

$$\text{Incidence rate} = \frac{\text{Number of new cases of disease or injury during specified period}}{\text{Time each person was observed (totaled for all persons)}}$$

Incidence rate for this study = 3 cases / 236 person-days

= 0.0127 cases per person-day

Make it's easier to understand:

0.0127 cases per person-day x 1000 = 12.7 cases per 1000 person-days

Interpret: There is 12.7 recurrent cases of MI per 1000 persons per day

Now, we would like to know "per year" (time units convertion)

Back to before x 1000 persons

= (0.0127 cases per person-day) x (365 person-day/1 person-year)

= 4.6 cases person-years

Interpret: There is 4.6 recurrent cases of MI persons per day

- Typically, each person is observed from an established starting time until one of four "end points" is reached
    - End-point means:
        - onset of disease
        - death
        - migration out of the study ("lost to follow-up")
        - end of the study

• An incidence rate describes how quickly disease occurs in a population. It is based on person-time, so it has some advantages over an incidence proportion. Because person-time is calculated for each subject, it can accommodate persons coming into and leaving the study.

• Person-time has one important drawback. Person-time assumes that the probability of disease during the study period is constant, so that 10 persons followed for one year equals one person followed for 10 years. Because the risk of many chronic diseases increases with age, this assumption is often not valid.

Example of incidence proportion and incidence rate[1]

In a study in the United States of America, the incidence rate of stroke was measured in 118,539 women who were 30–55 years of age and free from coronary heart disease, stroke and cancer in 1976

Table 2.4. Relationship between cigarette smoking and incidence rate of stroke in a cohort of 118 539 women[13]

Smoking category	Number of cases of stroke	Person-years of observation (over 8 years)	Stroke incidence rate (per 100 000) person-years)
Never smoked	70	395 594	17.7
Ex-smoker	65	232 712	27.9
Smoker	139	280 141	49.6
Total	274	908 447	30.2

Incidence proportion (cumulative incidence) for stroke over the eight-year follow-up

= 274 / 118539

= 0.0023

Make it easier to understand: x 1000 persons

= 0.0023 x 1000 persons = 2.3 stork cases per 1000 persons

[1] Bonita, R., Beaglehole, R., & Kjellström, T. (2006). *Basic epidemiology*. World Health Organization

Prevalence

Prevalence is a proportion of persons in a population who have a particular disease or attribute at a specified point in time or over a specified period of time. It includes all cases, both new and preexisting, in the population at the specified time.

$$\text{Prevalence} = \frac{\text{All new and pre-existing cases during a given time period}}{\text{Population during the same time period}} \times 10^n$$

The value of 10^n is usually 1 or 100 for common attributes. The value of 10^n might be 1,000, 100,000, or even 1,000,000 for rare attributes and for most diseases.

Point prevalence refers to the prevalence measured at a particular point in time. It is the proportion of persons with a particular disease or attribute on a particular date.
eg: Do you currently have asthma?

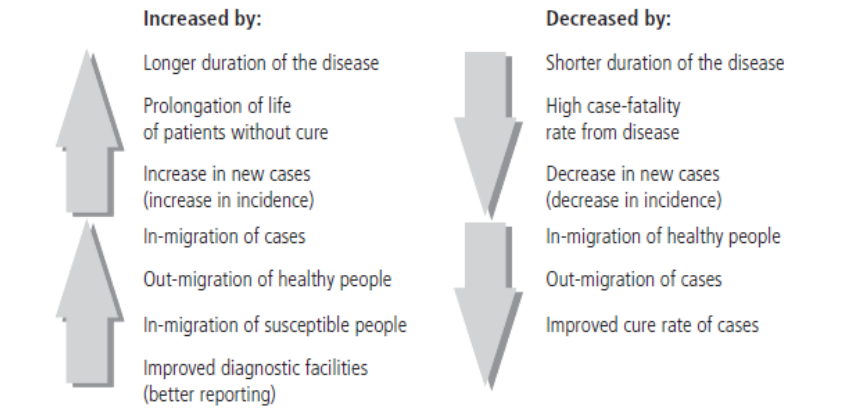
Period prevalence refers to prevalence measured over an interval of time. It is the proportion of persons with a particular disease or attribute at any time during the interval.
eg. Have you had asthma during last year?

Prevalence and incidence variation

A chronic, incurable disease, such as diabetes, can have a low incidence but high prevalence, because the disease is not very fatal.

A short-duration, curable disease, such as the common cold, can have a high incidence but low prevalence, because many people get a cold each year.

Figure 2.2. Factors influencing prevalence

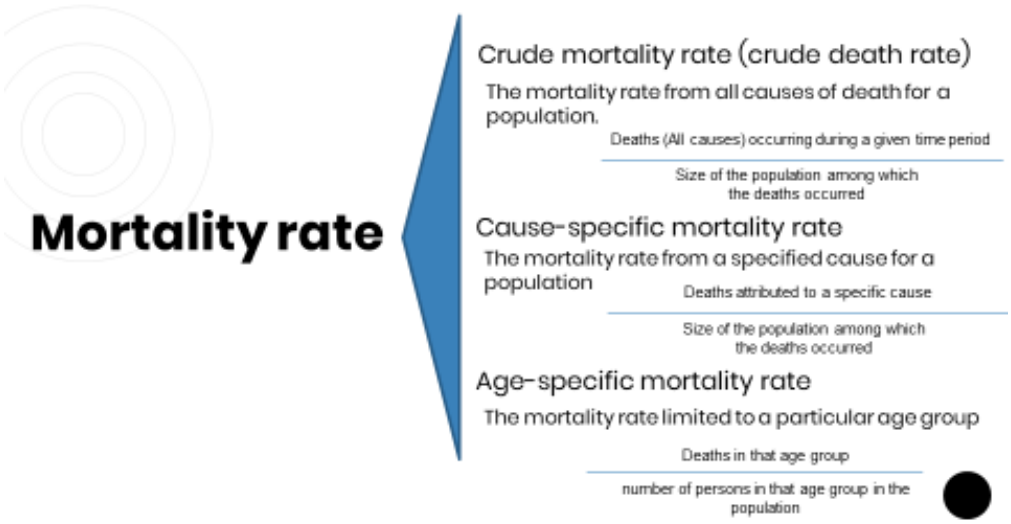

Increased by:
Longer duration of the disease
Prolongation of life of patients without cure
Increase in new cases (increase in incidence)
In-migration of cases
Out-migration of healthy people
In-migration of susceptible people
Improved diagnostic facilities (better reporting)
Decreased by:
Shorter duration of the disease
High case-fatality rate from disease
Decrease in new cases (decrease in incidence)
In-migration of healthy people
Out-migration of cases
Improved cure rate of cases

Measuring Mortality frequency

Mortality rate

Mortality rate is the frequency of occurrence of death in a defined population during a specified interval

$$\text{Mortality rate} = \frac{\text{Deaths occurring during a given time period}}{\text{Size of the population among which the deaths occurred}}$$


Mortality rate
Crude mortality rate (crude death rate)
The mortality rate from all causes of death for a population.
Deaths (All causes) occurring during a given time period
Size of the population among which the deaths occurred
Cause-specific mortality rate
The mortality rate from a specified cause for a population
Deaths attributed to a specific cause
Size of the population among which the deaths occurred
Age-specific mortality rate
The mortality rate limited to a particular age group
Deaths in that age group
number of persons in that age group in the population

Age-adjusted mortality rate (Age-standardize the rate)

Age-adjusted mortality rate means a mortality rate statistically modified to eliminate the effect of different age distributions in the different populations.

Note:

Mortality rates obviously increase with age, a higher mortality rate among one population than among another might simply reflect the fact that the first population is older than the second is.

Standardization Using an External Distribution

Distribution of the US Population

Age Group	Population (% of Total)
<5	18,300,000 (7%)
5-19	52,900,000 (22%)
20-44	98,100,000 (40%)
45-64	46,000,000 (19%)
>64	30,400,000 (12%)
Total	245,700,000 (100%)

Age-specific Mortality Rates in Florida

Age Group	Number of People	% of Total Pop.	Death Rate per 100,000
<5	850,000	7%	284
5-19	2,280,000	18%	57
20-44	4,410,000	36%	198
45-64	2,600,000	21%	815
>64	2,200,000	18%	4,425
Totals	**12,340,000**	**100%**	**1,069 (crude rate)**

Age-specific Mortality Rates in Alaska

Age Group	Number of People	% of Total Pop.	Death Rate per 100,000
<5	60,000	11%	274
5-19	130,000	25%	65
20-44	240,000	45%	188
45-64	80,000	15%	629
>64	20,000	4%	4,350
Totals	**530,000**	**100%**	**399 (crude rate)**

Standardization Using an External Distribution (US pop. Distribution)

Florida = 797 per 100,000 population

For the age group <5 years old: 0.07 x 284 =

For the age group 5 to 19 years: 0.22 x 57 =

For the age group 20 to 44 years: 0.40 x 198 =

For the age group 45 to 64 years: 0.19 x 815 =

For the age group greater than 64 year: 0.12 x 4,425 =

Alaska = 750 per 100,000 population

For the age group <5 years old: 0.07 x 274 =

For the age group 5 to 19 years: 0.22 x 65 =

For the age group 20 to 44 years: 0.40 x 188 =

For the age group 45 to 64 years: 0.19 x 629 =

For the age group greater than 64 year: 0.12 x 4,350 =

Death-to-case ratio

is the number of deaths attributed to a particular disease during a specified time period divided by the number of new cases of that disease identified during the same time period.

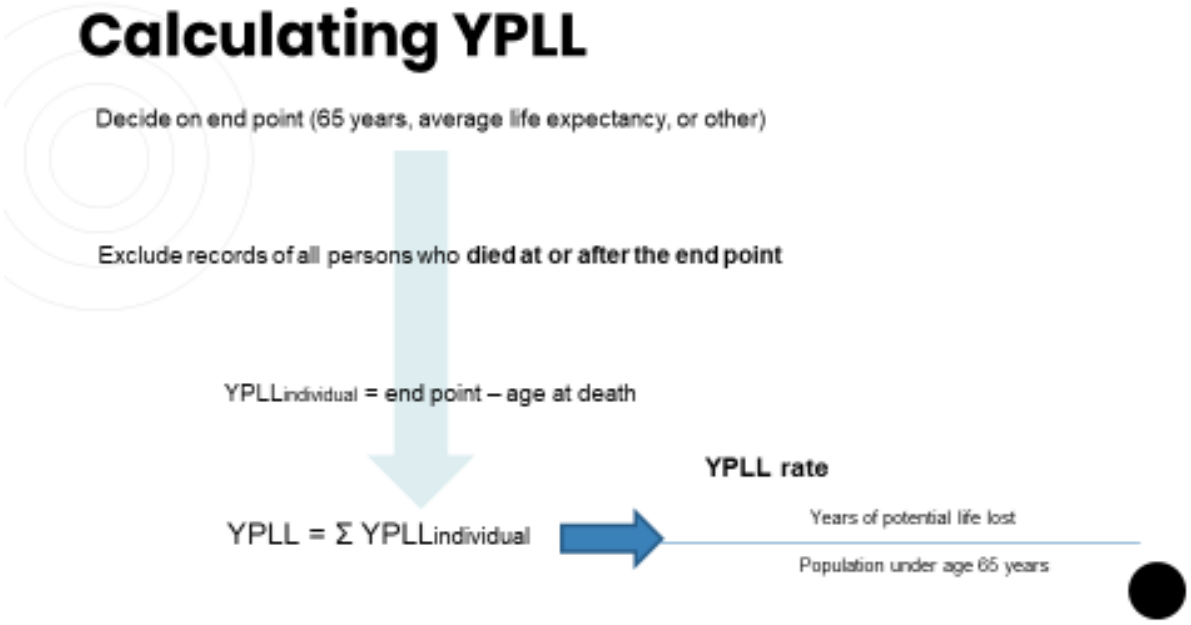
$$\textbf{Death-to-case ratio} = \frac{\text{Number of deaths attributed to a particular disease during specified period}}{\text{Number of new cases of the disease identified during the specified period}}$$

Years of potential life lost (YPLL)

Years of potential life lost is one measure of the impact of premature mortality on a population

Two most commonly used end points

- age 65 years
- average life expectancy


Calculating YPLL
Decide on end point (65 years, average life expectancy, or other)
Exclude records of all persons who died at or after the end point
YPLLindividual = end point – age at death
YPLL = Σ YPLLindividual
YPLL rate
Years of potential life lost
Population under age 65 years

Remark: Measuring of associations are included in session 4, session 5 and session 7

5300501 – Lecture # Session 4 Epidemiological study: Cross sectional study

Learning Objectives

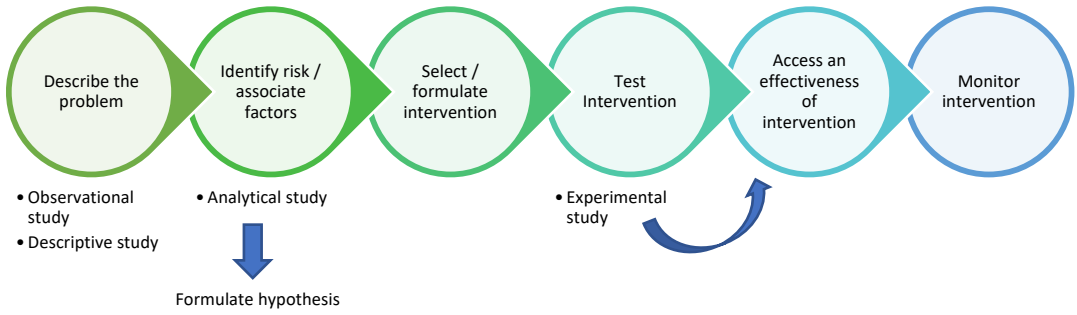
- Understand structure of cross-sectional study (descriptive and analytical)
- Identify exposure and health outcomes in cross-sectional study
- Construct 2 x 2 table for association measurement calculation
- Be able to calculate and interpret association measurement

Study design in epidemiology

"A study design is a specific plan or protocol for conducting the study, which allows the investigator to translate the conceptual hypothesis into an operational one"

By: Ahmed Mandil, MBChB, DrPH, University of Alexandria

Flow Diagram for public health research [1]



[1] Adapted from de Zoysa, I., Habicht, J. P., Pelto, G., & Martines, J. (1998). Research steps in the development and evaluation of public health interventions. Bulletin of the World Health Organization, 76(2), 127.

Cross-sectional study

Cross sectional studies are carried out at a single point (or we call "snapshot period") when we need to compare with other type of studies (figure xx). We usually perform cross sectional study to assess the prevalence of the outcome of any interest (may be disease) for a given population.

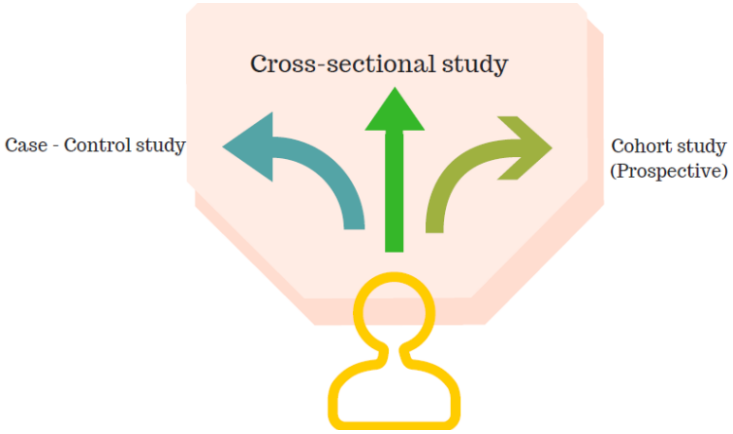


Figure 9 *Study design 1*

Cross-sectional studies are divided into descriptive and analytical cross-sectional study.

1. **Descriptive cross-sectional study** uses for describing the prevalence of a health outcome in a specified population.
    - Prevalence can be both types, **point prevalence** and **period prevalence**. Point prevalence explains prevalence of disease at specific point of time while period prevalence is considered time period (normally, we consider accumulate sufficient information on disease in a specific population)
        - Point prevalence refers to the prevalence measured at a particular point in time. It is the proportion of people with a particular disease or attribute on a particular date.
        - Period prevalence refers to prevalence measured over an interval of time (the "point in time" is broader). It is the proportion of people with a particular disease or attribute at any time during the interval.

Point prevalence

= Number of current cases (new and preexisting) at a specified point in time
Population at the same specified point in time

Period prevalence
= Number of current cases (new and preexisting) over a specified period of time
Average or mid-interval population

Example of Point and Period prevalence [1]

Point Prevalence

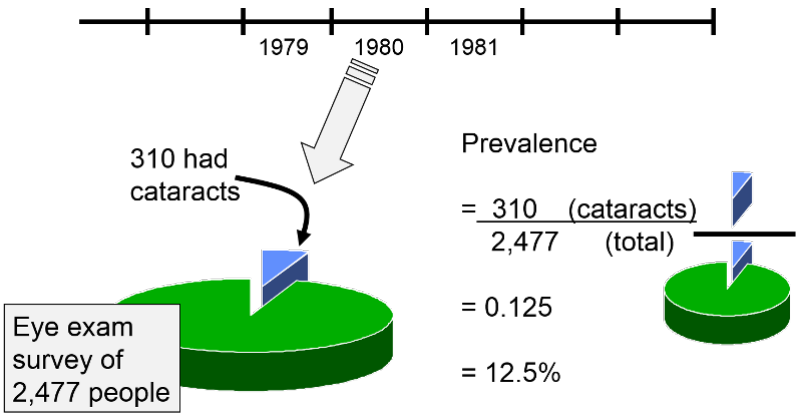
E.g. The percentage of a class reporting symptoms of seasonal allergies during the first week in May 2016.

- We include all the students who have symptoms before the class start and still have during the first week. Students whom developed symptoms during that week are also counted.

Period prevalence

E,g. Suppose that 2,477 residents of Framingham, MA were examined in order to establish the proportion of the population who had cataracts. It would have taken 2-3 years to conduct all of the eye exams, and when examinations were done, the prevalence over this observation period have included people who had acquired cataracts earlier assuming that they still lived in that populations. It also included new cases, i.e. those who had developed cataracts during the 2-3 year period when the eye exams were conducted. So, it could just be thought of as a wide "point in time".

During 1980 the Framingham Het Study examined 2,477 subjects for cataracts and found 310 of them had cataracts . Therefore, the prevalence was 310/2,477 = 0.125.



[1] LaMorte, W. Measures of Disease Frequency. Boston University School of Public Health Available at https://bit.ly/2LuEL5a

- Normally, we describe characteristic by 4W (What / Who / Where / When – we keep Why for analytical study)
  - "What" means what disease and severity that we are going to describe.
  - "Who" means specific / defined population
  - "Where" means place (in Bangkok, Thailand, South-East Asia, Asia, or World)
  - "When" means a snap shot period. (When we take a photo)

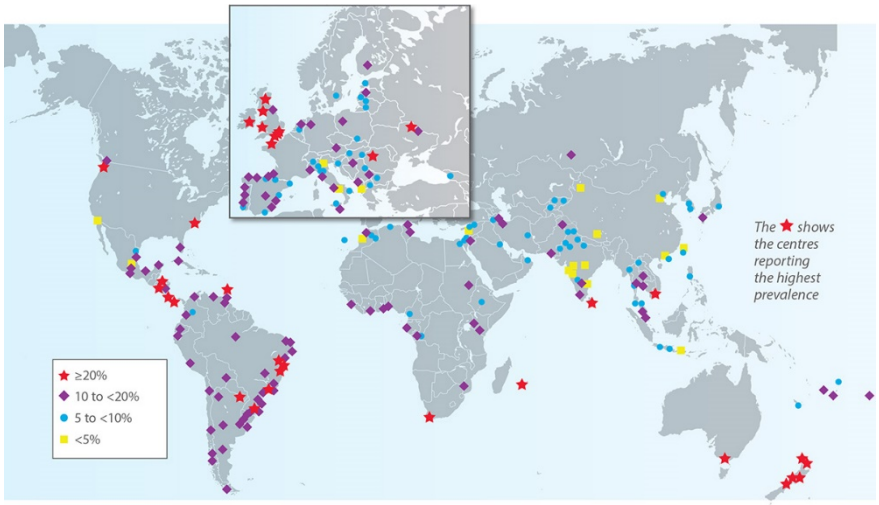


Figure 1: Prevalence of asthma symptoms among 13-14 year olds (ISAAC).

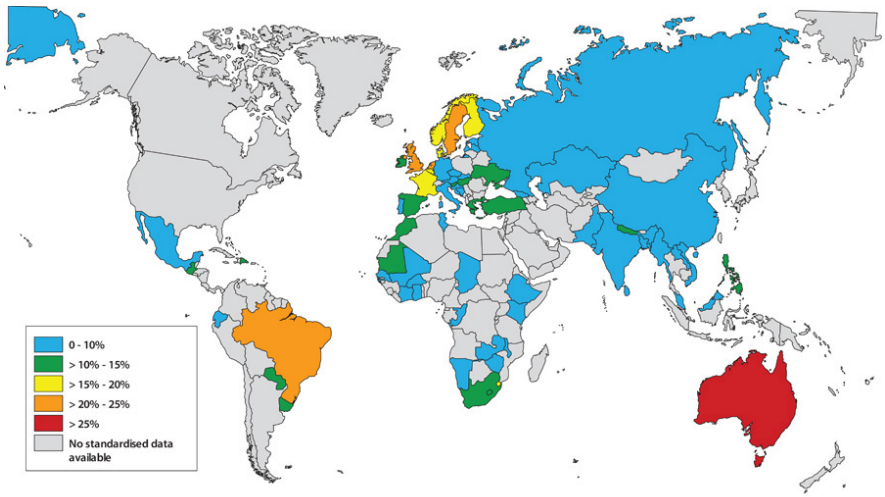


Figure 3: Prevalence of symptoms of asthma in the past 12 months among persons aged 18 to 45 years in 70 countries, World Health Survey 2002-2003.

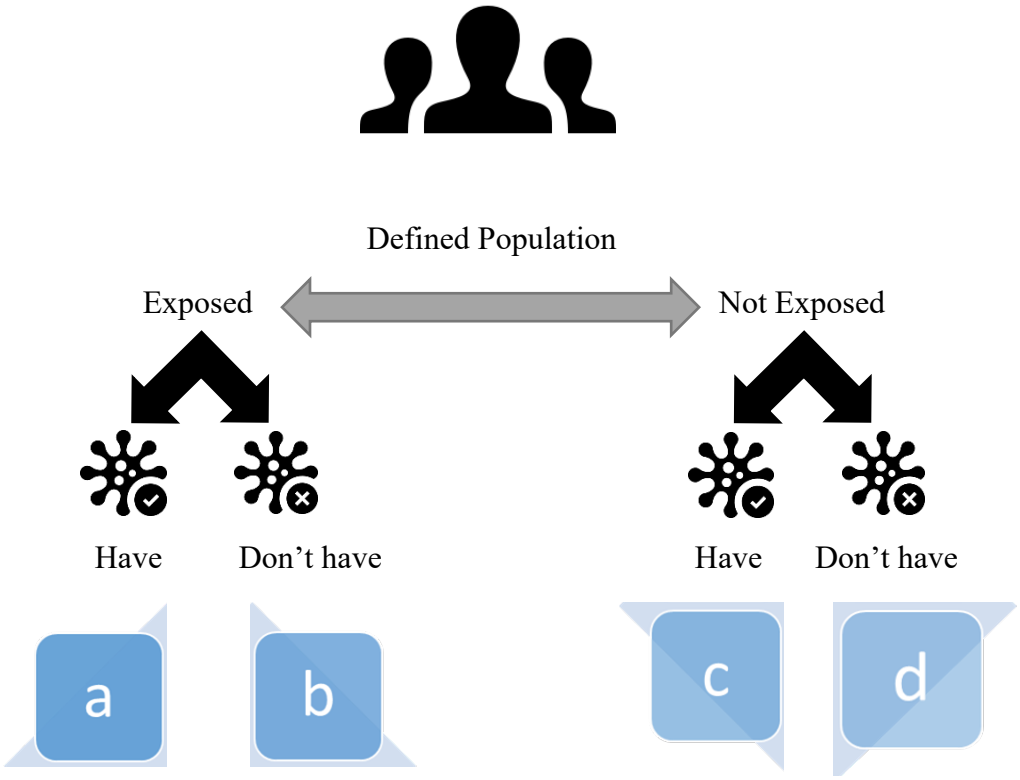
Figure 10 *Example of point prevalence and period prevalence*[1]

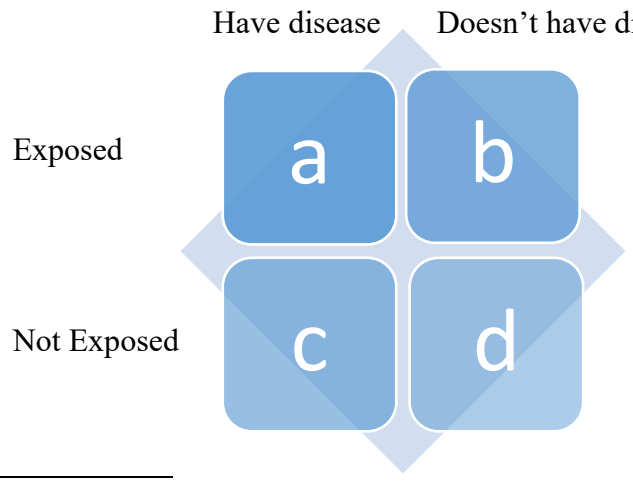
[1] The Global Asthma Report 2014. Available online: http://www.globalasthmareport.org/burden/burden.php

2. **Analytical cross-sectional studies** are used to identify the possible factors associated with health outcome. We will be trying to compare the difference of health outcome among exposure and non-exposure group.

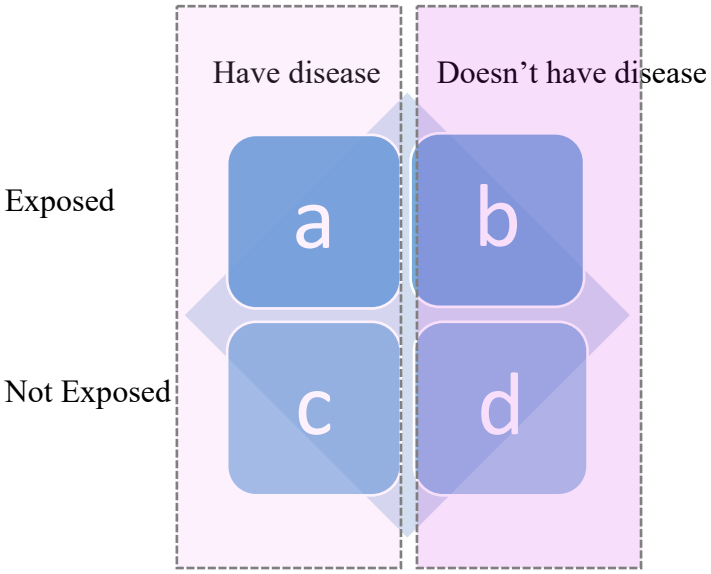
Structure of analytical cross-sectional data[1]

In analytical cross-sectional studies both prevalence of exposure and health outcome can be obtained. The purpose of this study is to compare the health outcome differences between exposed and unexposed.





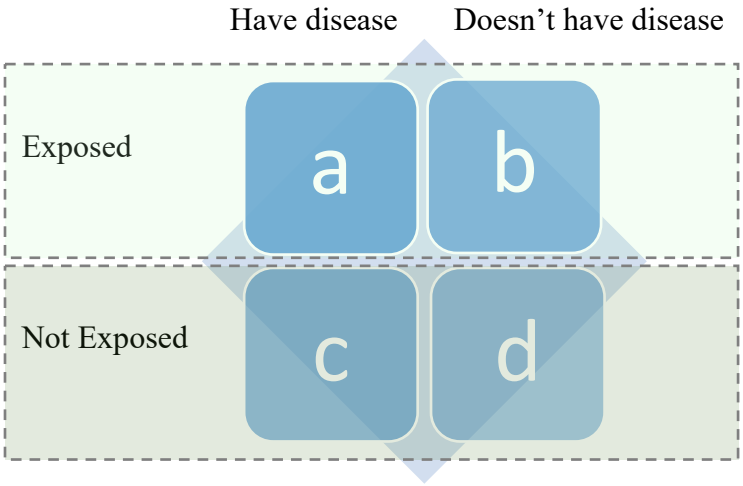
[1] Apply from John Hopskin



Prevalence of exposure among disease = a / (a + c)

Prevalence of exposure among no disease = b / (b + d)

Try!! Prevalence of not exposure among disease?



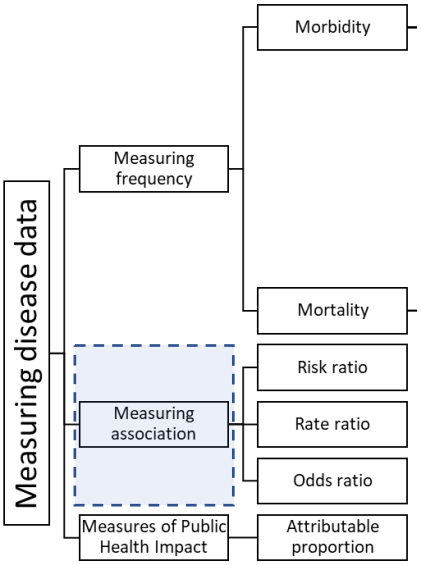
Prevalence of disease among exposure = a / (a + b)

Prevalence of disease among no exposure = c / (c + d)

Try!! Prevalence of no disease among no exposure?

Measuring an association in Cross-sectional Study

In order to identify the association between exposure and health outcomes prevalence ratio (a nickname of "risk ratio") and odd ratio are applied for the analysis. Besides the proportion calculation, we would like to compare those proportion for seeing the association.



Odds ratio (OR) (In cross-sectional study, sometimes it is called prevalence odd ratio because prevalent cases are included)

The prevalence odds ratio (POR) is calculated as same as the odds ratio. Odds ratio (OR) is a measure of association between an exposure and an outcome. The OR represents the odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of that exposure.

$$\text{OR (POR)} = \frac{a \times d}{b \times c}$$

	Have disease	**No have disease**
Exposure	a	b
Non-exposure	c	d

Note:

The probability that an event will occur is the fraction of times you expect to see that event in many trials. Probabilities always range between 0 and 1.

The odds are defined as the probability that the event will occur divided by the probability that the event will not occur. The odds of an event represent the ratio of the (probability that the event will occur) / (probability that the event will not occur).

Odds of event = Y / (1-Y)

*Y = Probability that an event will happen

	Have disease (Case)	**No have disease (Control)**
Exposure	a	b
Non-exposure	c	d

$$\text{Odds E cases} = \frac{\text{Cases with exposure}}{\text{Cases without exposure}} = \frac{a}{c}$$

$$\text{Odds E control} = \frac{\text{Control with exposure}}{\text{Control without exposure}} = \frac{b}{d}$$

To get an odd ratio of CASE:

$$\text{Odds E cases} = \frac{\text{Cases with exposure}}{\text{Cases without exposure}} = \frac{a}{c}$$

$$\text{Odds E control} = \frac{\text{Control with exposure}}{\text{Control without exposure}} = \frac{b}{d}$$

$$\text{OR} = \frac{a \times d}{b \times c}$$

Interpretation:

OR=1 Exposure does not affect odds of outcome

OR>1 Exposure associated with higher odds of outcome (Risk factor)

OR<1 Exposure associated with lower odds of outcome (Protective factor)

Prevalence ratio (a nickname of "Risk ratio")

The prevalence ratio (PR) is analogous to the risk ratio (RR) of cohort studies.. The populations are fixed as denominators for both ratios in a cohort study – fixed at the beginning of the study and fixed at the point or period of time in case-control study.

	Have disease	**No have disease**	
Exposure	a	b	a+b
Non-exposure	c	d	c+d

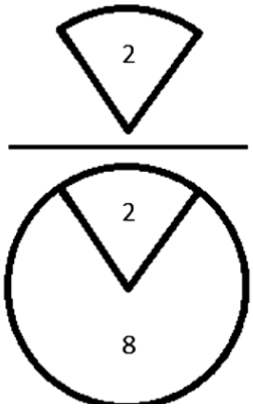
Prevalence ratio = $\frac{\text{Prevalence of disease among exposed group}}{\text{Prevalence of disease among non-exposure group}}$

= $\frac{a / (a + b)}{c / (c + d)}$

For chronic disease studies or studies of long-lasting risk factors, POR is the preferred measure of association in cross-sectional studies.

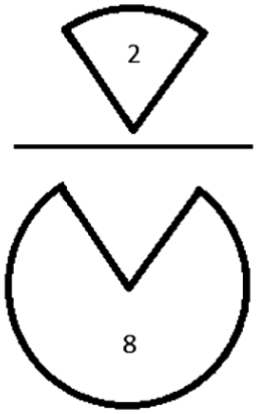
For acute disease studies, PR is the preferred measure of association. If the prevalence of disease is low, i.e. 10% or less in exposed and unexposed populations, POR = PR. Since cross-sectional studies are particularly useful for investigating chronic diseases, where the onset of disease is difficult to determine or for studying long lasting risk factors (such as smoking, hypertension and high fat diets), **the prevalence odds ratio** will generally be the preferred measure of association.

The difference between odd and risk [1]



The number of patients with the risk factor and the outcome in the numerator, and all patients with the risk factor in the denominator

(Risk = 2 ⁄ (8 + 2) = 2 ⁄ 10 = 20%).



The number of patients with the risk factor and the outcome in the numerator and the number of patients with the risk factor without the outcome in the denominator

(Odds = 2 ⁄ 8 = 1 ⁄ 4 = 0.25)

When do we have to use cross-sectional studies?

Descriptive cross-sectional study is an important method to evaluate the proportion of a population with disease or with risk factors for disease. They are widely used to estimate the occurrence of risk factors in specific population which may be characterized by age, sex, race or socioeconomic status (SES).

Examples of descriptive cross-sectional study:

- National Health Survey in each country
- Opinion polls and political polls
- Surveillance of behavioral risk factors

Descriptive cross-sectional studies are useful for planning or administering preventive or health care services, surveillance programs and surveys and polls

[1] Wilber, S. T., & Fu, R. (2010). Risk ratios and odds ratios for common events in cross-sectional and cohort Studies. Academic Emergency Medicine, 17(6), 649-651.

Analytical cross-sectional studies are useful for establishing preliminary evidence for a causal relationship even though we cannot conclude the relationship. At least, we get some idea for continuing future study. These studies are also useful for examining the association between exposure and disease onset for chronic diseases where researchers lack information on time of onset. However, interpretation requires caution regarding potential association of duration of disease with exposure status.

Advantages of cross-sectional studies

- Inexpensive and takes little time to conduct
- Can estimate prevalence of outcome of interest as sample is usually taken from the whole population
- Many outcomes and risk factors can be assessed
- Useful for public health planning, understanding disease's etiology and for the generation of hypotheses
- No loss to follow-up

Disadvantages of cross-sectional studies

- Difficult to make causal inference
- Only a snapshot: the situation may provide differing results if another timeframe had been chosen
- Cannot conclude causal relationship between exposure and health outcomes

Bias in cross-sectional study

By design

1. Prevalence-incidence bias (also called Neyman bias) especially in the case of long-lasting diseases. Any risk factor that results in death will be under-represented among those with the disease.

2. Antecedent-consequent bias, similar to the chicken and egg question (i.e. "which came first?"). This bias occurs when it cannot be determined that exposure preceded disease, since both are ascertained at the same time (unlike cohort studies or clinical trials). Antecedent-consequent bias does not affect cohort studies because subjects in cohort studies are selected for study as they are disease-free. Exposure is actually observed to precede disease only in a cohort design, including randomized trials.

3. Survival bias may be minimized if information can be obtained on exposures that clearly preceded the first symptoms of a chronic disease such as arthritis, diabetes or chronic bronchitis. This depends on access to data (medical records) before the onset of a chronic disease. In addition, it may be necessary to have historical records on an individual's exposure status prior to these first medical visits.

By study participants[1]

1. Selection bias

Ideally, conducting study in population reduces selection bias. But, it is not possible. So, it needs to select a sample of subjects from a pool (study population) that can be very large. Obtaining a representative sample is crucial but not easy task. Ideally, one should have the full list of eligible subjects and the sample is picked using simple random sampling with every subject standing an equal chance of being included. Multistage sampling and systematic sampling, with caution, can sometimes achieve the goal of having a representative sample.

2. Information bias

In cross-sectional studies, information on risk factors and health conditions (outcomes), as well as other factors is often obtained at the same time-point. Adopting standardised and validated methods and using objective measures can help to avoid information biases. Self-reporting (answering a questionnaire or responding agreeing to be interviewed) is frequently used to collect a substantial proportion of the information in these studies.

[1] Yu. I., Tse. S. Workshop 6 — Sources of bias in cross-sectional studies; summary on sources of bias for different study designs. Hong Kong Med J. 18 (3). 2012

5300501 – Lecture # Session 5 Epidemiological study: Case Control study

Learning Objectives

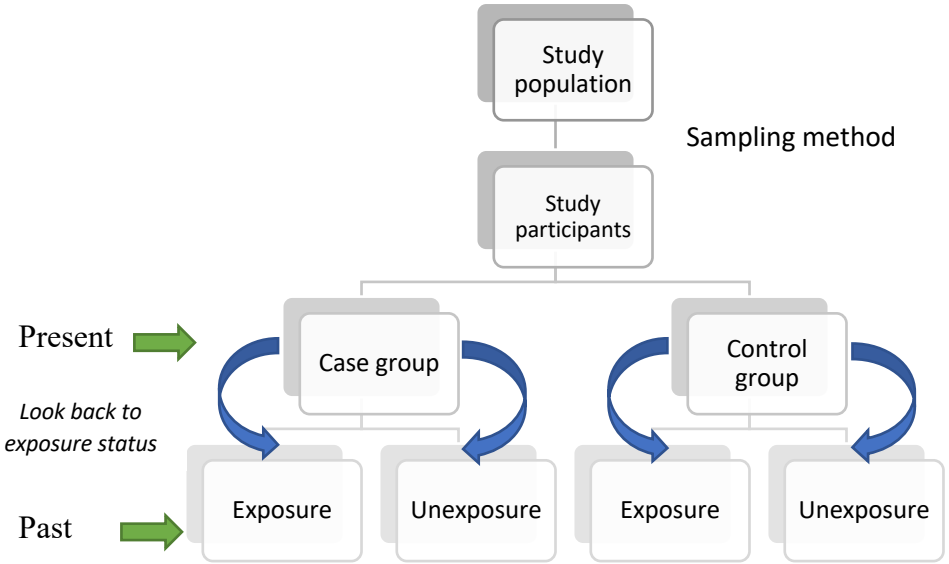
- Understand structure of case-control study design
- Know source of cases and controls
- Identify association measurement & be able to calculate and interpret

Case-Control Study

Case-control studies are used to determine an association between an exposure and a specific health outcome. This study design can proceed from effect (health consequences – "outcome") to cause (we call "exposure"). Case-control is normally used for *rare disease and rare health outcomes*.

Case-control study is separated into 2 types

1. Case – control
2. Matched case-control

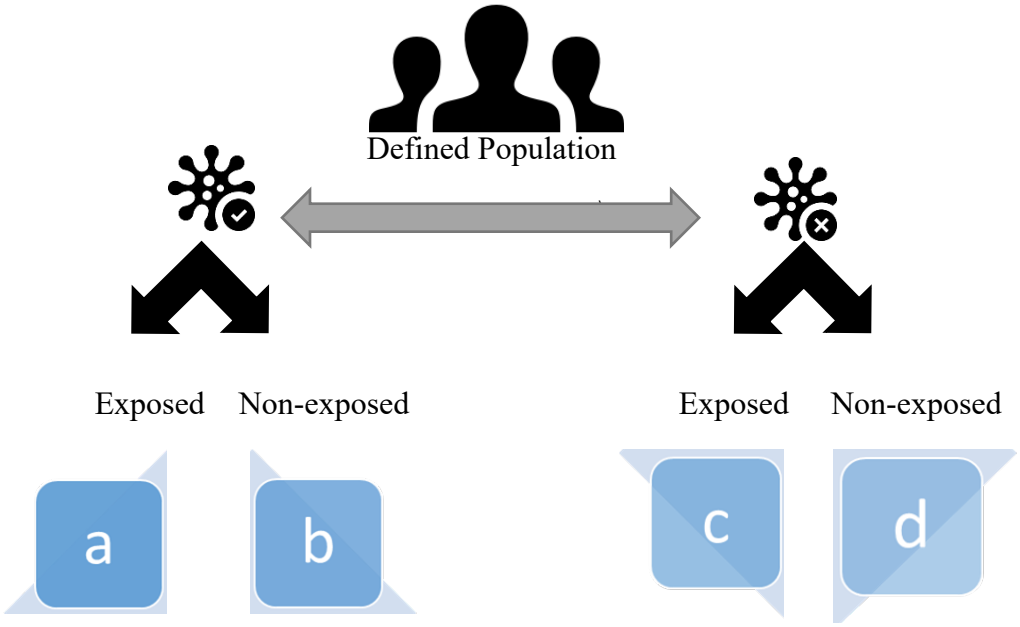


Unlike cohort or cross-sectional studies, subjects in case-control studies are selected because they have the health outcome of interest (cases). Selection is not based on exposure status. Controls, persons who are free of the health outcome under study, are randomly selected from the population out of which the cases arose. The case-control study aims to achieve the same goals (comparison of exposed and unexposed) as a cohort study but does so more efficiently, by the use of sampling.

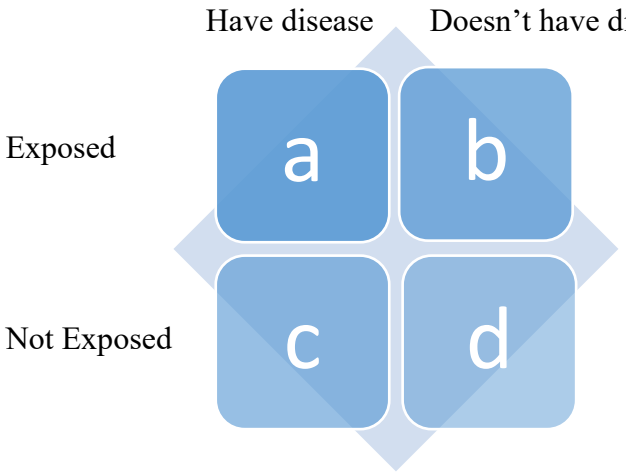
After cases and controls have been identified, the investigator determines the proportion of cases and the proportion of controls that have been exposed to the exposure of interest. Thus, the denominators obtained in a case-control study do not represent the total number of exposed and non-exposed persons in the source population

Structure of case-control study[1]

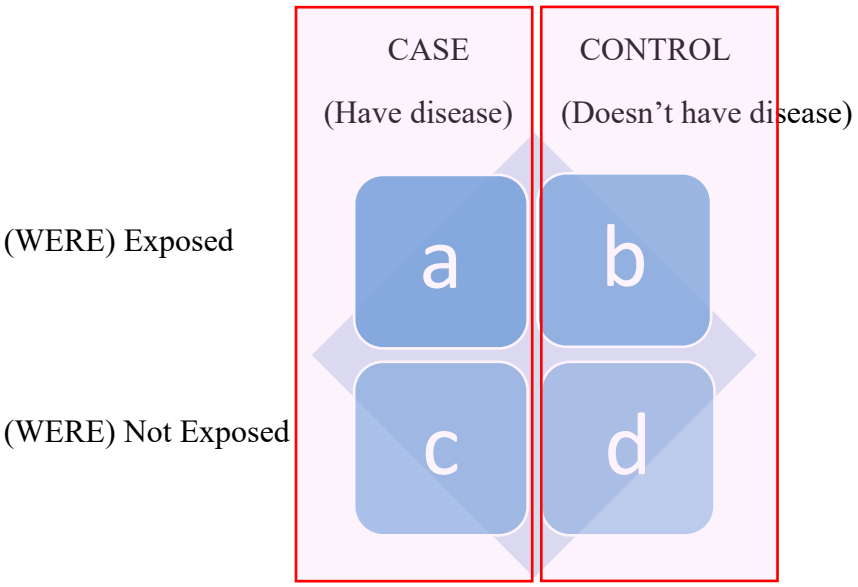
In case-control studies, health outcomes are obtained before identifying exposure. The purpose is for investigation of risk factors among health outcomes.



2 x 2 table is constructed in the same format as analytical cross-sectional study as following;



[1] Apply from John Hopskin



Proportion of exposed among case and control (sometimes we convert to percentage)

Proportion of exposed among case = a / (a + c)

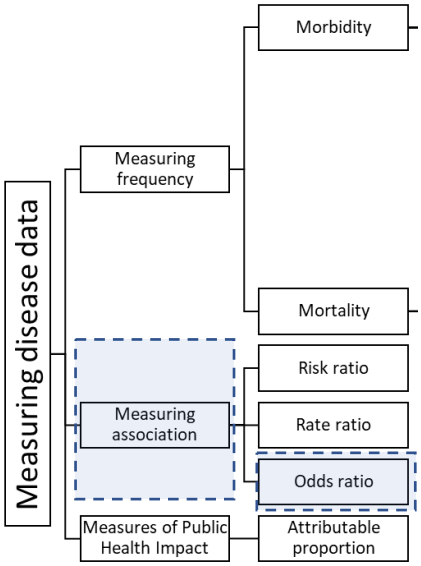
Proportion of exposed among control = b / (b + d)

Note:

Calculation is the same as prevalence of exposure among disease in analytical study because prevalence is a the proportion.

Prevalence of exposure among disease = a / (a + c)

Measuring an association in Case-control Study



In case-control study, odd ratio can be obtained. Measures of incidence in case-control studies are not possible. In case-control studies the proportion of cases in the entire population-at-risk is unknown, therefore one cannot measure incidence of the health outcome or disease. The controls are representative of the population-at-risk, but are only a sample of that population, therefore the denominator for a risk measure, the population-at-risk, is unknown. We decide on the number of diseased people (cases) and non-diseased people

(controls) when we design our study, so the ratios of controls to cases is not biologically or substantively meaningful.

Odds ratio (OR)

Odds ratio (OR) is used for identify risk of exposure over the health outcomes. The OR represents the odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of that exposure.

$$OR = \frac{a \times d}{b \times c}$$

	Have disease	**No have disease**
Exposure	a	b
Non-exposure	c	d

Odds Ratio as an Estimate of Risk Ratio[1]

Provided that the disease is uncommon (say <10%), this sampling approach gives an odds ratio that is a reasonably good estimate of the risk ratio. Very rare outcomes (e.g., in the tables above) will give odds ratios that are extremely close to what the risk ratio would be. However, as the outcomes of interest become more common, the odds ratio gives estimates that are increasing more extreme than the risk ratio would have been. By more extreme, I mean that odds ratios that are greater than 1 will be larger than the corresponding risk ratio, and odds ratios that are less than 1 will be smaller than the corresponding risk ratio.

[1] LaMorte WW. Measuring Association in Case-Control Studies. Boston University School of Public Health. Available online: https://bit.ly/2M8Z8oK

If the outcome is common (let's say >10%), the odds ratio exaggerates the estimated strength of association. And, OR cannot predict Risk ratio (RR).

	Disease	Non-disease	Cumulative incidence
Exposed	60	108	60/(60+108) = 35.7%
Non-exposed	45	341	45/(45+341) = 11.7%

Risk ratio = $\frac{60 / (60 + 108)}{45 / (45 + 341)}$ = 3.06

OR = $\frac{60 / 108}{45 / 341}$ = 4.21

The interactive feature below allows you to simultaneously compute both the risk ratio and the odds ratio in a hypothetical cohort study. In general, the odds ratio will be close in value to the risk ratio when the outcome of interest is rare, but the odds ratio will tend to become more extreme than the risk ratio as the outcome becomes increasingly common. You can use this widget to observe how the odds and probabilities change as you make the outcome more common. For example, enter numbers in the cells that would be consistent with a rare outcome and compare the OR and the RR. Then increase the frequency of the outcome by doubling, quadrupling, etc. the number of events in the first column without changing the second column.

If the outcome is uncommon (let's say < 10%), we can predict Risk Ratios from Odds Ratios.

Observed Data

	Diseased	No Disease	Total
Exposed	7	1000	1007
Non-exposed	6	5634	5640
Total	13	6634	6647

Incidence in exposed= (7 / 1007) = 0.00695134

Incidence in non-exposed= (6 / 5640) = 0.00106383

Risk Ratio = 6.5343

Odds of disease in exposed= (7 / 1000) = 0.00700000

Odds of disease in exposed= (6 / 5634) =0.00106496

Odds Ratio = 6.5730

Types of case-control studies

Case-control studies can be categorized into different groups based on when the cases develop the health outcome and based on how controls are sampled. Some case-control studies use prevalent cases while other case-control studies use incident cases. There are also different ways that cases can be identified, such as using population-based cases or hospital-based cases.

Types of cases used in case control studies

Prevalent cases are all persons who were existing cases of the health outcome or disease during the observation period. These studies yield a prevalence odds ratio, which will be influenced by the incidence rate and survival or migration out of the prevalence pool of cases, and thus does not estimate the rate ratio. Case control studies can also use incident cases, which are persons who newly develop the health outcome or disease during the observation period. Recall that prevalence is influenced by both incidence and duration. Researchers that study causes of disease typically prefer incident cases because they are usually interested in factors that lead up to the development of disease rather than factors that affect duration.

Selecting controls

Selection of controls is usually the most difficult part of conducting a case-control study. We will discuss 3 possible ways to select controls:

1. Base or case-base sampling

2. Cumulative density or survivor sampling

3. Incidence density or risk set sampling

Base sampling or case-base sampling

This sampling involves using controls selected from the source population such that every person has the same chance of being included as a control. This type of sampling only works with a previously defined cohort. In these case-control studies,

the odds ratio provides a valid estimate of the risk ratio without assuming that the disease is rare in the source population.

Cumulative density sampling or survivor sampling

When controls are sampled from those people who remained free of the health outcome at the end of follow-up then we call the sampling cumulative density sampling or survivor sampling. Controls cannot ever have the outcome (become cases) when using this type of sampling. In these case-control studies, the odds ratio estimates the rate ratio only if the health outcome is rare, i.e. if the proportion of those with the health outcome among each exposure group is less than 10% (requires the rare disease assumption).

Incidence density sampling or risk set sampling

When cases are incident cases and when controls are selected from the at-risk source population at the same time as cases occur (controls must be eligible to become a case if the health outcome develops in the control at a later time during the period of observation) then we call this type of sampling incidence density sampling or risk set sampling. The control series provides an estimate of the proportion of the total person-time for exposed and unexposed cohorts in the source population. In these case-control studies, the odds ratio estimates the rate ratio of cohort studies, without assuming that the disease is rare in the source population.

Note that it is possible, albeit rare, that a control selected at a later time point could become a case during the remaining time that the study is running. This differs from case-control studies that use cumulative density sampling or survivor sampling, which select their controls after the conclusion of the study from among those individuals remaining at risk.

Selecting controls in a risk set sampling or incidence density sampling manner provides two advantages:

1. A direct estimate of the rate ratio is possible.

2. The estimates are not biased by differential loss to follow up among the exposed vs. unexposed controls.

For example, if a large number of smokers left the source population after a certain time point, they would not be available for selection at the end of the study – when controls would be selected in a study that uses cumulative density sampling or survivor sampling. This would give the investigators biased information regarding the level of exposure among the controls over the course of the study.

Source populations for case-control studies

Source populations can be restricted to a population of particular interest, e.g. postmenopausal women at risk of breast cancer. This restriction makes it easier to control for extraneous confounders in the population. Controls should represent the restricted source population from which cases arise, not all non-cases in the total population. The cases in the study do not have to include all cases in the total population.

Sources of cases

- Cases diagnosed in a hospital or clinic
- Cases entered into a disease registry, e.g. cancer, birth defects, deaths
- Cases identified through mass screening, e.g. hypertensives, diabetics
- Cases identified through a prior cohort study, e.g. lung cancers in an occupational asbestos cohort

Sources of controls

- Population controls are non-cases sampled from the source population giving rise to cases. This is the most desirable method for selecting controls. Sampling randomly from census block groups, or a registry such as the Department of Motor Vehicles (of adults who are able to drive) are examples of ways to find and recruit population-based controls.
- Neighborhood or friend controls are appropriate for selection as controls if these individuals would be included as cases if they developed the health outcome of interest. It is not appropriate to select neighbors or friends as controls if they share the exposure of interest.
- Hospital controls - There are certain problems with hospital controls in that they may not be from the same source population from which the cases arose. Hospital controls may not be representative of the exposure prevalence in the source population of cases, e.g. there may be a higher prevalence of smokers in hospitals. Hospital controls also may have diseases resulting from the exposure of interest, e.g. the exposure (smoking) is related to the disease of interest (cancer) and to heart and lung diseases from which the controls may be suffering.
- Controls with another disease - However if the study is on lung cancer, for example, it is essential to exclude cancers known or suspected to be related to the study exposure of interest. These controls also share some of the same problems as hospital controls.

Advantages of case-control studies

Case-control studies are the most efficient design for rare diseases and require a much smaller study sample than cohort studies. Additionally, investigators can avoid the logistical challenges of following a large sample over time. Thus, case-control studies also allow more intensive evaluation of exposures of cases and controls. Case-control studies that use incidence density sampling or risk set sampling yield a valid estimate of the rate ratio derived from a cohort study if incident cases are studied and controls are sampled from the risk set of the source population. If properly performed (i.e. appropriate sampling), case-control studies provide information that mirrors what could be learned from a cohort study, usually at considerably less cost and time.

Disadvantages of case-control studies

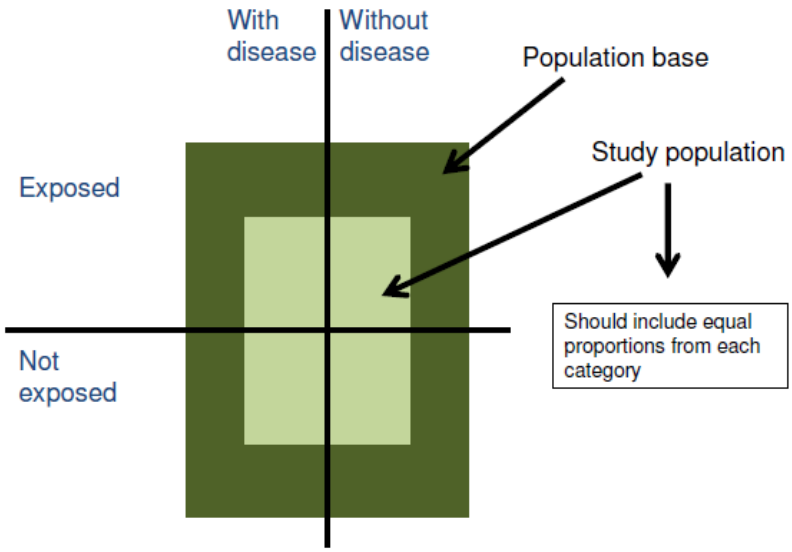
Case-control studies do not yield an estimate of rate or risk, as the denominator of these measures is not defined. Case-control studies may be subject to recall bias if exposure is measured by interviews and if recall of exposure differs between cases and controls. However, investigators may be able to avoid this problem if historical records are available to assess exposure. Choosing an appropriate source population is also difficult and may contribute to selection bias. Case-control studies are not an efficient means for studying rare exposures (less than 10% of controls are exposed) because very large numbers of cases and controls are needed to detect the effects of rare exposures

Bias in case-control study

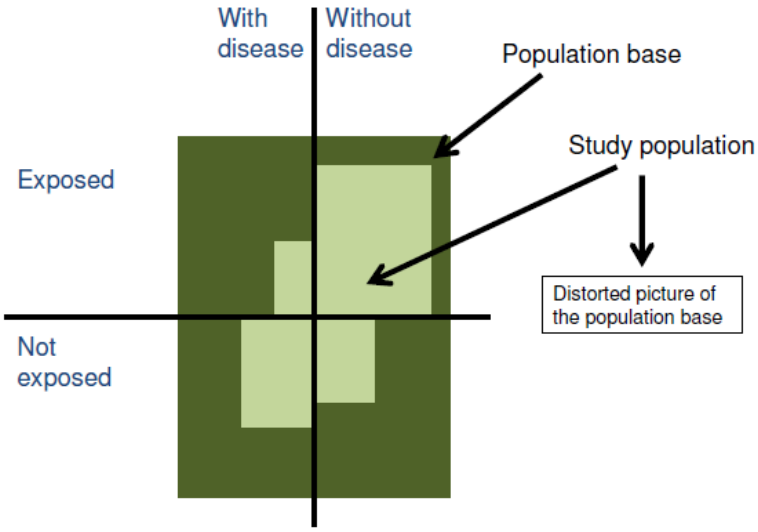
Selection bias

Selection bias occurs when a systemic error in the ascertainment of cases or controls in case-control studies. If exposure status is differentially distributed between cases and controls, leading to a distortion of the exposure-disease association.

Ideally, the proportion of sample selection should be;



In the real sampling technique, it will be like this;



Information Bias

• Often referred to as measurement bias

• Occurs due to poor measurement (classification) of study variables (exposure)

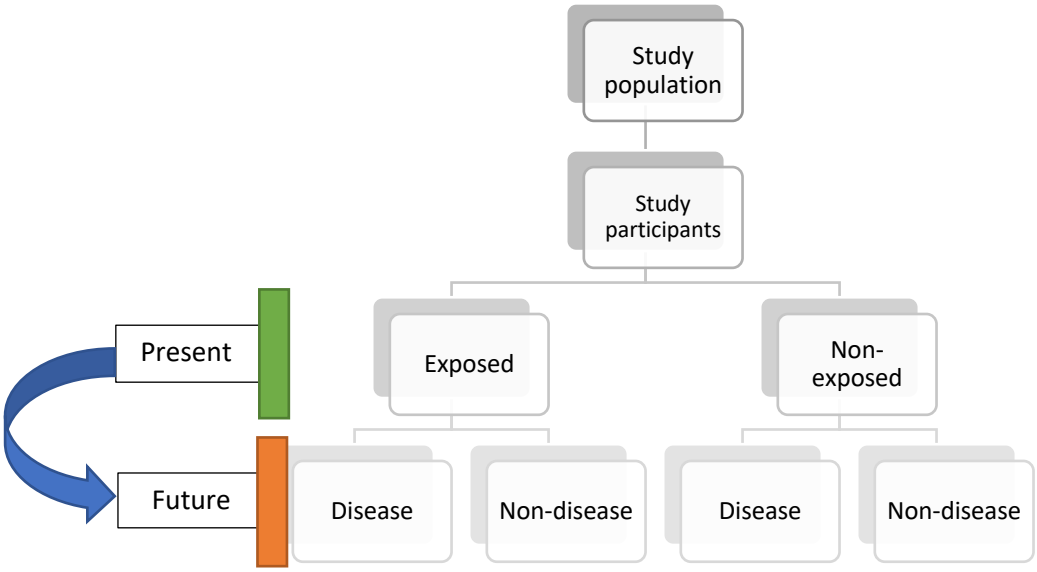
5300501 – Lecture # Session 7 Epidemiological study: Cohort Study

Learning Objectives

- Understand structure of cohort study design
- Understand Close and open cohort
- Characterize exposure and outcome in cohort study
- Able to identify measurement in cohort studies

What's cohort study?

A cohort study is a type of epidemiological study in which a group of people with a common characteristic is followed over time to find how many reach a certain health outcome of interest (disease, condition, event, death, or a change in health status or behavior). Cohort studies compare an exposed group of individuals to an unexposed (or less exposed) group of individuals to determine if the outcome of interest is associated with exposure.



A general idea for cohort study are;

- Investigator has to be defined 2 or more than 2 groups of exposure who are "free of disease" BUT "at risk"
- Some cohorts begin with a group of population that is HETEROGENOUS with respect to exposure. Subgroups are made by defining exposure later.

Type of cohort study I

There are two types of cohort studies: prospective and retrospective (or historical) cohorts. Prospective studies follow a cohort into the future for a health outcome, while retrospective studies trace the cohort back in time for exposure information after the outcome has occurred. Both types of cohort studies are also referred to as longitudinal or follow-up studies.

Comparison of prospective and retrospective cohorts

	Baseline	**Follow-up**
Prospective cohort	Assessed at beginning of study	Followed into the future for outcome
Retrospective cohort	Assessed at some point in the past via historical records	Outcome has already occurred and is assessed via historical records

Note:

Prospective studies

Exposure are measured for the purpose of research *before* the outcomes events occur.

Retrospective studies

Exposure are measured for the purpose of research after the outcomes events occur. (Both exposure and outcomes had been happened before study.)

Type of cohort study II

Cohort studies also classified into 2 groups by population in the study.

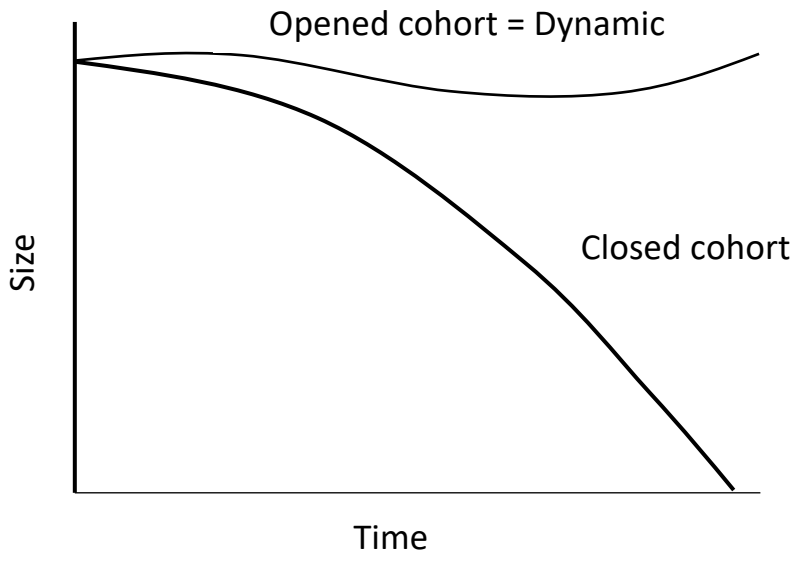


Figure 11 Size of opened and closed cohort

Closed cohort

A closed cohort is one with fixed membership. Once the cohort is defined by enrolling subjects and follow up begins, no one can be added. The number of subjects may decline because of death or loss to follow up, but no additional subjects are added. As a result, closed cohorts always get smaller over time.

- Once participants are member, always member
- No new members can enter
  - Eg. People born in 1985 / Students registered at school in 2018
- The membership defining event can be defined on several time scales
  - A cohort born in a particular year (calendar time)
  - A cohort formed at entry into military (an event)
  - A cohort of 50 year olds accrued over several years
  - A cohort of steel workers with 5 years experience
- On some time scales the starting time may not be the same for all members

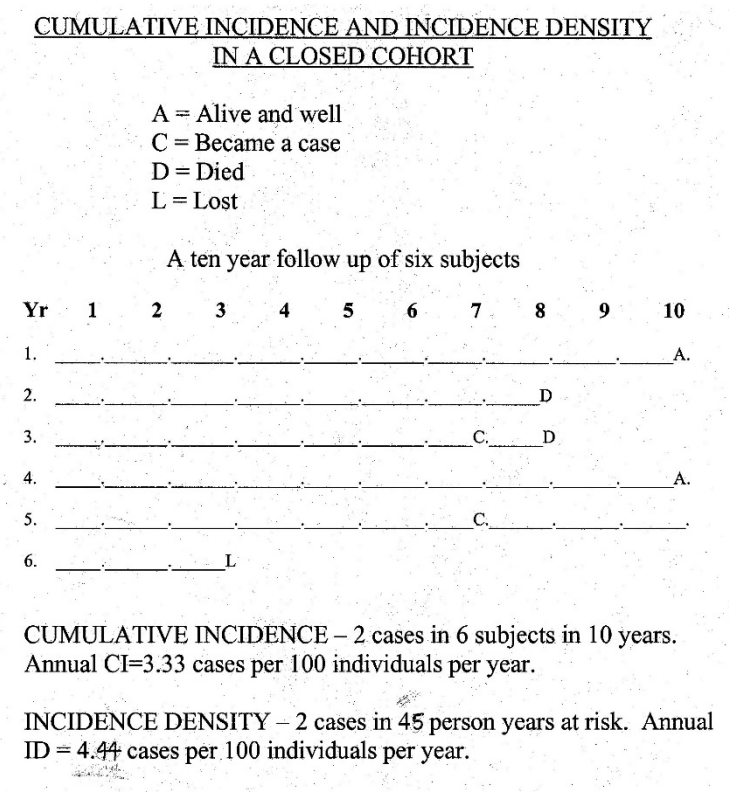

CUMULATIVE INCIDENCE AND INCIDENCE DENSITY
IN A CLOSED COHORT
A = Alive and well
C = Became a case
D = Died
L = Lost
A ten year follow up of six subjects
Yr 1 2 3 4 5 6 7 8 9 10
1. ____.____.____.____.____.____.____.____.____.____A.
2. ____.____.____.____.____.____.____.____D
3. ____.____.____.____.____.____.____C.____D
4. ____.____.____.____.____.____.____.____.____.____A.
5. ____.____.____.____.____.____.____C.____.____.____.
6. ____.____.____L
CUMULATIVE INCIDENCE – 2 cases in 6 subjects in 10 years.
Annual CI=3.33 cases per 100 individuals per year.
INCIDENCE DENSITY – 2 cases in 45 person years at risk. Annual
ID = 4.44 cases per 100 individuals per year.

Figure 12 Structure of Closed Cohort

Opened cohort

An open cohort is dynamic, meaning that members can leave or be added over time. Given the example of a state cancer registry, subjects are continually added when they are diagnosed with cancer, so new subjects are continually added. Subjects can also leave the cohort by moving to a new state or dying.

- Membership can vary over time
- There is no list of members, only **person-time** accrued from potentially changing roster of individuals
    - Current residents of Bangkok
    - In-patients of one ward at a hospital
- Each person contributes small or large amounts of person-time.
- One person can come and go overtime
    - Eg. Mr.A was a member of a health plan A for 1 year then switched to another health plan B, 6 months later he rejoined the plan A again.

      Mr.A contributed person-time to the open cohort defined as being a current member of the health plan A only during period in which he was a member of plan A.

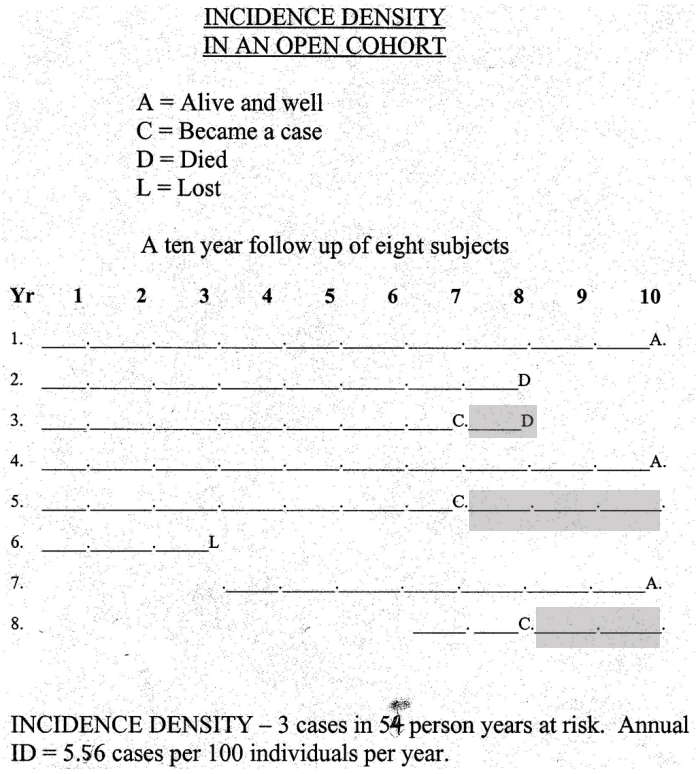

INCIDENCE DENSITY
IN AN OPEN COHORT
A = Alive and well
C = Became a case
D = Died
L = Lost
A ten year follow up of eight subjects
Yr 1 2 3 4 5 6 7 8 9 10
1. A.
2. D
3. C. D
4. A.
5. C.
6. L
7. A.
8. C.
INCIDENCE DENSITY – 3 cases in 54 person years at risk. Annual ID = 5.56 cases per 100 individuals per year.

Figure 13 Structure of Opened Cohort

The great majority of cohort studies are conducted in closed (or fixed) cohorts, because it is more difficult to establish eligibility and track people in an open cohort, since they can enter and leave at any time. This problem becomes greater as the size of the cohort gets larger and/or the study continues for a longer period of time.

Person-time

Person-time is an estimate of the actual time-at risk in years, months, or days that all persons contributed to a study. In certain studies people are followed for different lengths of time, as some will remain disease-free longer than others. A subject is eligible to contribute person-time to the study only so long as that person remains disease-free and, therefore, still at risk of developing the disease of interest. By knowing the number of new cases of disease and the person-time-at-risk contributed to the study, an investigator can calculate the incidence rate of the disease, or how quickly people are acquiring the disease.

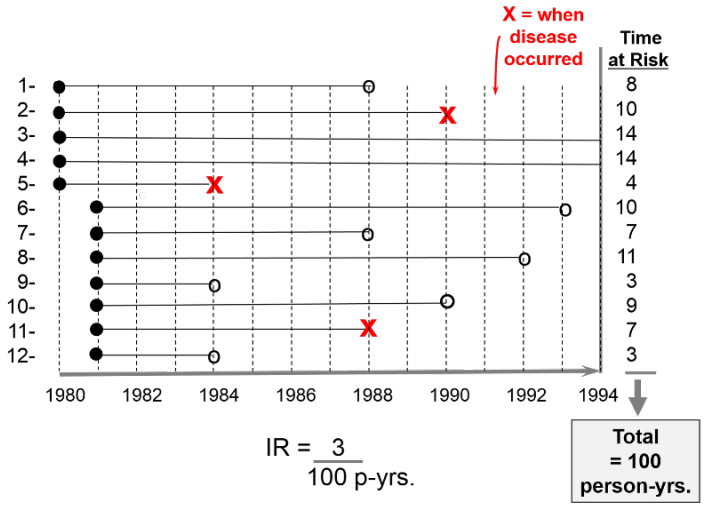


Figure 14 *An example of person-time calculation*

How to conduct cohort

Establishing the cohort

The investigator controls the selection of the cohort. The investigator may choose a cohort based on age, location, exposure to a certain working environment, or some other common characteristic. Cohorts may be selected on the basis of exposures known at baseline, e.g. smokers vs. nonsmokers. Alternatively, cohorts may be divided into exposure categories once baseline measurements of a defined population are made.

For example, the Framingham Cardiovascular Disease Study (CVD) used baseline measurements to divide the population into categories of CVD risk factors.

For instance, an investigator wants to study whether exposure to the noise of military aircraft engine is a risk factor for loss of hearing. The cohort this investigator would want to establish should be composed of two groups of military personnel: one exposed to engine aircraft noise (the group under study) and the other unexposed to engine aircraft noise (a comparison group). The cohort this investigation should be composed of two groups of military personnel: one exposed to the noise of engine aircraft (the group under study) and the other unexposed to the noise of engine aircraft (a comparison group). The unexposed group should be the representative of the exposed group for all factors except exposure.

The cohort at baseline

After the cohort of study subjects is established, their individual exposures of interest are identified at baseline (through interviews, questionnaires, bioassays, medical records, etc.). Subjects with the outcome of interest at baseline are excluded. Therefore, all members of the cohort are at risk of developing the outcome at the beginning of observation.

Following the last example, anyone in the cohort of military personnel with a specified hearing loss at baseline would be excluded from the cohort and would not be followed.

Following the cohort

The cohort is then followed over time for new occurrences of the outcome of interest, in the above example, hearing loss. In a prospective, or concurrent, cohort study baseline exposure is assessed at the beginning of the study and the cohort is followed into the future. In a retrospective, or historical cohort study, baseline exposure is assessed at some point in the past through historical records, e.g. health records for a cohort of factory workers may provide exposure and outcome information up to the present.

Evaluation of the results

During the follow-up period the investigator counts the number of subjects who develop the outcome of interest. This count is the numerator for a calculation of risk, also referred to as cumulative incidence or incidence proportion. The number of persons at risk at baseline is the denominator.

Measuring of disease frequency in Cohort Study

- In closed cohort, we can measure,
  - Cumulative incidence (Incidence proportion)
  - Incidence rate (Incidence density)
  - Odds (incidence Odds)

Incidence proportion or risk (Cumulative incidence)

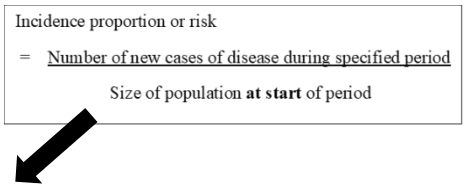
Incidence proportion or risk = Number of new cases of disease during specified period Size of population **at start** of period

Incidence proportion is the proportion of an initially disease-free population that develops disease (die) **at specified period of time.**

PROPORTION - the numerator is included in the denominator

Incidence proportion is a proportion because the persons who develop disease are all included in the entire population.

Synonyms for incidence proportion
• Attack rate
• Risk
• Probability of developing disease
• Cumulative incidence


Incidence proportion or risk
= Number of new cases of disease during specified period
Size of population at start of period

Population at start of period (Denominator)

- Should be limited to the "population at risk" for developing disease

For example : the new cases is ovaries cancer patients, the denominator should be restricted to women.

During disease outbreak, the term "**attack rate"** is often used as a synonym for risk. It is the risk of getting the disease during a specified period, such as the duration of an outbreak.

- Overall attack rate is the total number of new cases divided by the total population.
- A food-specific attack rate is the number of persons who ate a specified food and became ill divided by the total number of persons who ate that food

Incidence rate (person-time) (Incidence density)

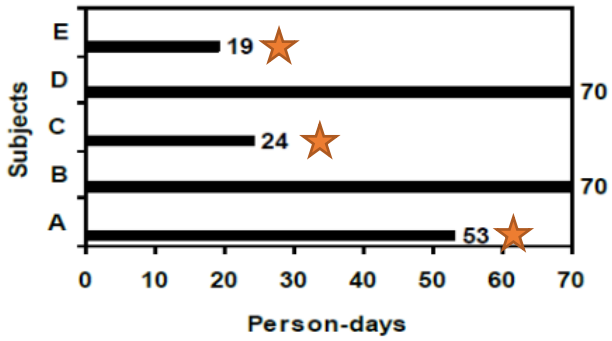
Incidence rate = Number of new cases of disease or injury during specified period Time each person was observed (total for all persons)

RATE: The denominator is the sum of follow-up time contributed by the people at risk of the event.

Person-time

- A person-time rate is generally calculated from a long-term cohort follow-up study.
- Person-time is an estimate of the actual time-at-risk (in years (normally use), months, or days) – that all participants contributed to a study.

Example [1]



★ Developed second MI

Suppose one researcher is conducting a study of the rate of second myocardial infarction (MI). She follow 5 patients which have first MI up to 10 weeks. The results is showed above.

X-axis presents number of days for each person contribute to the study.

[1] Alexander, L. K., Lopes, B., Ricchetti-Masterson, K., & Yeatts, K. B. Calculating Person-Time.

Y-axis presents 5 patients who were follow up in this study.

From the graph, we can calculate person-days (unit is days) for 5 patients who contribute to the study.

Time contributed by each subject:

Subject A: 53 days

Subject B: 70 days

Subject C: 24 days

Subject D: 70 days

Subject E: 19 days

Total person-days in the study: 53+70+24+70+19=236 person-days

All ★ means patients who developed second MI. So, there are 3 patients who have recurrent of MI during 10 week (70 days) of follow-up.

Then, we convert to incidence rate.

Incidence rate

$$= \frac{\text{Number of new cases of disease or injury during specified period}}{\text{Time each person was observed (totaled for all persons)}}$$

Incidence rate for this study = 3 cases / 236 person-days

= 0.0127 cases per person-day

Make it's easier to understand:

0.0127 cases per person-day x 1000 = 12.7 cases per 1000 person-days

Interpret: There is 12.7 recurrent cases of MI per 1000 persons per day

Now, we would like to know "per year" (time units convertion)

Back to before x 1000 persons

= (0.0127 cases per person-day) x (365 person-day/1 person-year)

= 4.6 cases person-years

Interpret: There is 4.6 recurrent cases of MI persons per day

- Typically, each person is observed from an established starting time until one of four "end points" is reached

- End-point means:
  - onset of disease
  - death
  - migration out of the study ("lost to follow-up")
  - end of the study

• An incidence rate describes how quickly disease occurs in a population. It is based on person-time, so it has some advantages over an incidence proportion. Because person-time is calculated for each subject, it can accommodate persons coming into and leaving the study.

• Person-time has one important drawback. Person-time assumes that the probability of disease during the study period is constant, so that 10 persons followed for one year equals one person followed for 10 years. Because the risk of many chronic diseases increases with age, this assumption is often not valid.

Example of incidence proportion and incidence rate[1]

In a study in the United States of America, the incidence rate of stroke was measured in 118,539 women who were 30–55 years of age and free from coronary heart disease, stroke and cancer in 1976

Table 2.4. Relationship between cigarette smoking and incidence rate of stroke in a cohort of 118 539 women[13]

Smoking category	Number of cases of stroke	Person-years of observation (over 8 years)	Stroke incidence rate (per 100 000) person-years)
Never smoked	70	395 594	17.7
Ex-smoker	65	232 712	27.9
Smoker	139	280 141	49.6
Total	274	908 447	30.2

Incidence proportion (cumulative incidence) for stroke over the eight-year follow-up

= 274 / 118539

= 0.0023

Make it easier to understand: x 1000 persons

= 0.0023 x 1000 persons = 2.3 stork cases per 1000 persons

Odds (Incidence odds) *(remark: more explanation in cross-sectional session)*

[1] Bonita, R., Beaglehole, R., & Kjellström, T. (2006). *Basic epidemiology*. World Health Organization

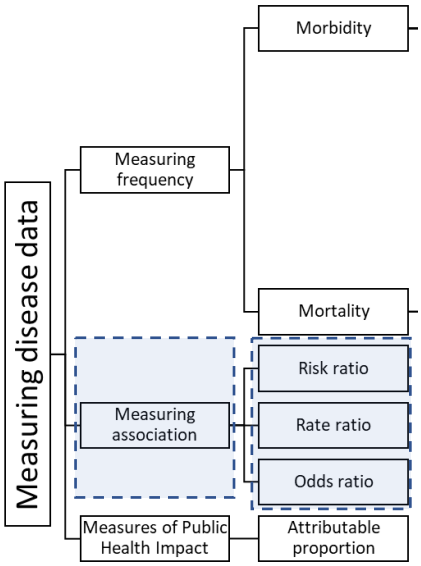
Odds of event = Y / (1-Y)

*Y = Probability that an event will happen

	Have disease	**No have disease**
Exposure	a	b
Non-exposure	c	d

Odds E cases = $\frac{\text{Cases with exposure}}{\text{Cases without exposure}} = \frac{a}{c}$

Measuring an association in Cohort Study



Cohort studies are longitudinal studies where an exposed and an unexposed group (or less exposed group) are followed forward in time to find the incidence of the outcome of interest (e.g. disease, death, or change in health). *Two measures of incidence are risks and rates.* Risks and rates can be further manipulated to provide additional information on the effects of the exposure of interest, such as risk ratios, rate ratios, attributable risks (risk or rate differences) and attributable risk percent.

Risk and Rate calculation

Risk is defined as the number of new cases divided by the total population-at-risk at the beginning of the follow-up period. An individual's risk of developing the outcome of interest is measured.

A rate is the number of new cases of a health outcome divided by the total person-time-at-risk for the population. Person-time is calculated by the sum total of time all individuals remain in the study without developing the outcome of interest (the total amount of time that the study members are at risk of developing the outcome of interest). Person-time can be measured in days, months, or years, depending on the unit of time that is relevant to the study. A rate measures the rapidity of health outcome occurrence in the population.

$$\textbf{Risk} = \frac{\text{Number of new cases}}{\text{Total of individuals at risk}}$$

$$\textbf{Rate} = \frac{\text{Number of new cases}}{\text{Total person-time at risk}}$$

Risk ratios (Relative risk)

When risks are computed in a study, the risk ratio is the measure that compares the Risk exposed to the Risk unexposed. The risk ratio is defined as the risk in the exposed cohort (the index group) divided by the risk in the unexposed cohort (the reference group). A risk ratio may vary from zero to infinity.

	Have disease	**No have disease**
Exposure	a	b
Non-exposure	c	d

$$\text{Risk ratio} = \frac{\text{Risk of exposed group}}{\text{Risk of unexposed group}}$$

$$\text{Risk ratio} = \frac{[a / (a+b)] / [\, c/ (c+d)]}{\text{Risk of unexposed group}}$$

Rate ratio

When rates are computed in a study, the rate is the measure that compares the Rate exposed to Rate unexposed . The rate ratio is defined as rate of health outcome occurrence in the exposed cohort index group) divided by the rate of occurrence in the unexposed or less-exposed cohort (reference group)

$$\text{Rate ratio} = \frac{\text{Rate of exposed group}}{\text{Rate of unexposed group}}$$

Interpretation of Risk and Rate

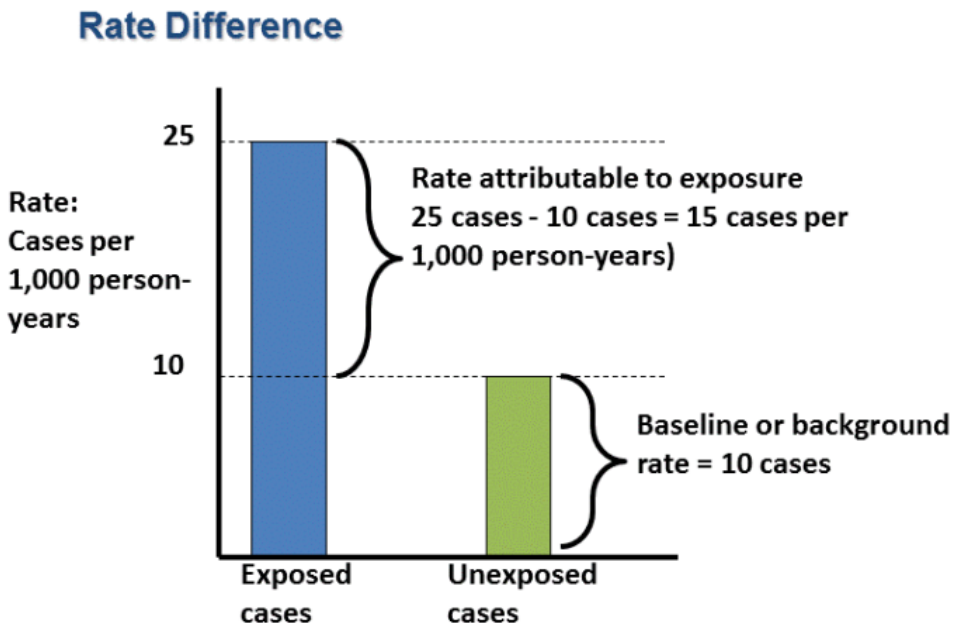
Risk ratio and Rate ratio	Exposure
<1	Exposure is protective
=1	Exposure is neither preventive nor harmful
>1	Exposure is harmful

The farther away the risk ratio or rate ratio is from the null value of one, the greater the effect of exposure is on the study group. This is shown in the following diagram.

Risk and Rate differences

Rate differences

In order to find the absolute effect of an exposure a health outcome the attributable rate (AR), or rate difference, must be computed. The term attributable rate is same as rate difference (RD). The attributable rate is the excess rate among the exposed population attributed to exposure. It is defined as the rate in the exposed minus the rate in the unexposed. The rate difference can also be reported as a percent (see figure below).



Adapted from Gordis L. Epidemiology, p. 156

Rate Difference = Rate of exposure – Rate of unexposed

Session 8 Individual evaluation report (10%)

Example of Assignment

Assignment Instruction

From the annual conference of our college, please answer the following questions;

Part 1 Contemporary Issues in Public Health from Prof. Rifat Atun (9.30 - 10.00)

1. Summarize main idea from the lecture
2. In your opinion, how to integrate epidemiology into contemporary Issues in Public Health ?

Part 2 Symposium Contemporary Issues in Public Health Enhancing sustainable health system for universal health coverage in Asia and the world (10.30 - 12.00)

1. Summarize (or compare) universal health coverage (UHC) of Japan and Thailand
2. In your opinion, what is a relationship between UHC and epidemiology data in country level? Please explain and give an example (Japan or Thailand).

Note: Assignment will be changed according to register semester

5300501 – Lecture # Session 9 Mid-term examination (Report – 15%)

From the case study, answer the following questions

Question 1: As a public health practitioner in Surin, what are the major concerns raised by these 121 cases?

Question 2: Critique the press release.

2.1 What type of bias would be introduced by this press release?

2.2 How might the press release impact the outbreak investigation?

Question 3: "Suppose the outbreak was linked to infected people coming from Bangkok and spread quickly"

3.1 What type(s) of study would you use to investigate this hypothesis? Why?

3.2 What general types of information would you include in this investigation?

Note: Assignment will be changed according to register semester

Example

Chikungunya virus infects 121 in Surin

© Bangkok Post : PUBLISHED : 19 JAN 2019 AT 16:59

According to the World Health Organization, Chikungunya is characterised by an abrupt onset of fever, frequently accompanied by joint pain. The joint pain can be debilitating, but most cases last for only a few days. Most patients recover fully, but in some cases joint pain may persist for months, or even years.

An outbreak of Chikungunya disease, a virus spread by mosquitoes, has struck three villages in Surin, with 121 people falling ill.

Theeraphol Sakulrat, chief of Sikhoraphum district, said affected residents of the villages in tambon Jaraphat had sought treatment at clinics in the province. Many suffered from high fever and pain in their ankles.

Health officials found 121 residents were infected, he said after visiting patients in the tambon where the outbreak was reported.

Hoei Maithong, 75, said she suffered from a high fever, skin rashes and joint pain and sought treatment at a local hospital. Many neigbhours also had the same symptoms, she added.

Plang Tor-ngam, 47, a resident of tambon Jaraphat, said her family members had recently recovered from the disease. She said the outbreak struck her community after some infected residents who worked in Bangkok recently returned home. After that the disease spread and other residents were infected.

Somsak Boontor, chief of the Sikhoraphum district public health office, confirmed that the outbreak was linked to infected people coming from Bangkok and spread quickly. However, he did not give further details.

Local health officials and volunteers have been spraying insecticides, burning garbage and clearing areas with stagnant water to reduce the control mosquito breeding, he added.

5300501 – Lecture # Session 10 Critical paper evaluation / Group presentation

Activities in this session:

1. Students will be divided into a small group (3 – 5 students/group) for discussion.
2. Research article together with questions will be sent to students at least 1 week prior to presentation date.
3. Students will make discussion within their group to answer the questions.
4. Presentation (15 – 20 min) will be made on indicated date.
5. All students in class will be participated by giving comments or suggestions.
6. After presentation, revised PowerPoint has to be sent back to lecturer.

Example

Article 1: Gelaye, B., Okeiga, J., Ayantoye, I., Berhane, H. Y., Berhane, Y., & Williams, M. A. (2016). Association of suicidal ideation with poor sleep quality among Ethiopian adults. Sleep and Breathing, 20(4), 1319-1326.

According to the article, please answer the following questions;

1. What is study design was used to examine the association between Suicidal Ideation with Poor Sleep Quality among Ethiopian Adults? Why do authors select this study design? Do you agree with this study design?
2. What is the exposure(s) of interest? What classification of the exposure did the authors use? What category was used to define the referent group?
3. What's was the primary association (s) of interest? What measure(s) of association was calculated to examine this association?
4. Calculate crude odds ratio for the association between having poor sleep quality compared to good sleep quality and suicidal ideation
5. What are the limitations of this study? Do you have any idea to deal with each limitation?

Article 2: Azoulay, L., Filion, K. B., Platt, R. W., Dahl, M., Dormuth, C. R.,Clemens, K. K.,& Targownik, L. E. (2016). Association between incretin-based drugs and the risk of acute pancreatitis. *JAMA internal medicine*, *176*(10), 1464-1473.

According to the article, please answer the following questions;

1. What's study design was used to examine the association between Incretin-Based Drugs and Acute Pancreatitis? Why do authors select this study design? Do you agree with this study design?
2. What's the base population for this study?
3. How were the cases and controls defined and identified? Do you think that is appropriate?
4. Did the authors match on any factors in the design of this study? If yes, please list the matched factors.
5. What's the exposure(s) of interest? What classification of the exposure did the authors use? What category was used to define the referent group?
6. What's was the primary association (s) of interest? What measure(s) of association was calculated to examine this association?
7. Calculate crude odds ratio for the association between using incretin based drugs compared to ≥ 2 oral antidiabetic drugs and the occurrence of acute Pancreatitis.
8. List potential confounders described by the authors. Why did they select these confounders ? Do you agree with the authors?

Note: Assignment will be changed according to register semester

5300501 – Lecture # Session 11 Environmental and Occupational Epidemiology

Learning Objectives

- To understand a linkage between environmental exposure and health outcomes
- To be able to apply epidemiological concept with environmental exposure assessment
- To conceptualize study design based on environmental and occupational epidemiology

Environmental factors and burden of diseases[1]

***"Environmental risks"* account for a large fraction of the global burden of disease**

In 2012, 12.6 million people died as a result of living or working in an unhealthy environment, representing 23% of all deaths. When both death and disability were counted the fraction of the global burden of disease due to the environment was 22%. Among children under five years old, up to 26% of all deaths were possible to prevent provided that environmental risks would have removed. 68% of these attributable deaths and 56% of attributable DALYs could be estimated with evidence-based comparative risk assessment methods and the impacts of other environmental exposures were assessed through expert opinion.

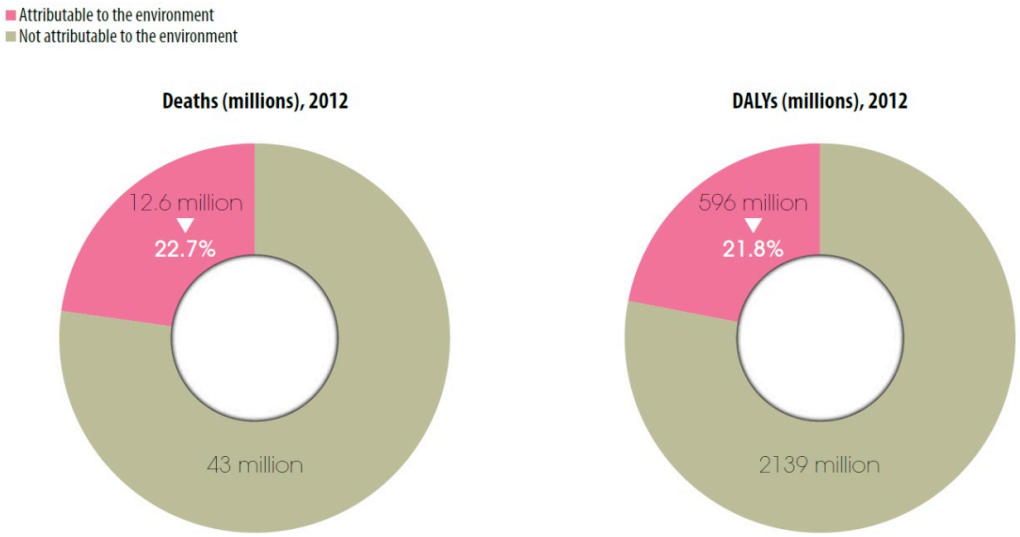


Figure Deaths and DALYs attributed to environment
(Ref https://www.who.int/gho/phe/en/)
**Combined measure of years of life lost due to mortality and years of life lost due to disability*

Five key results can be identified from the study on the impact of environmental risk factors on death and disease globally

1. Environmental risks account for a large fraction of the global burden of disease
A remarkable number of deaths are attributable to modifiable environmental risks, with 23% (95% CI: 13–34%) of all deaths, and 22% (95% CI: 13–32%) of the disease burden in DALYs – a combined measure of years of life lost due to mortality and years of life lost due to disability

[1] WHO, 2012 Derived from https://www.who.int/gho/phe/en/

2. Environmental impacts on health are uneven across life course and gender

The health of children under five, and to a lesser extent up to 10, and that of adults between 50 and 75 years are most affected group by the environment. In children, the environment's contribution to infectious and parasitic diseases, neonatal and nutritional diseases and injuries are very much prominent. In older adults, the fraction of noncommunicable diseases (NCDs) caused by the environment becomes more important and that for injuries remains constant but significant. Men are slightly more affected by the environment than women, with 22.8% versus 20.6% of all DALYs being attributable to the environment respectively (see Figure ES4). Women bear higher exposures to traditional environmental risks to health, such as exposure to smoke while cooking with solid fuels or carrying water from community sources. However, men are more exposed to occupational risks to health. Men may also be exposed to increased risks of injuries due to their patterns of occupational and recreational activities.

3. Low- and middle-income countries bear the greatest share of environmental disease

The highest fraction of deaths and disease, that could be tackled through environmental improvements is in the low- and middle-income countries

4. Total environmental deaths are unchanged since 2002, but show a strong shift to noncommunicable diseases

During the last decade it was seen a shift away from infectious, parasitic and nutritional diseases to NCDs, not only in terms of the environmental fraction but also the total burden. This shift is mainly due to a global decline of infectious disease rates and a reduction in the environmental risks causing infectious diseases i.e. a higher share of people having access to safe water and sanitation and a lower share of households using solid fuels for cooking. In terms of the total disease burden, NCDs have increased globally. While the highest number of deaths per capita attributable to the environment occurs in sub-Saharan Africa, primarily from infectious diseases, other regions now have higher rates of NCDs attributable to the environment.

The diseases with the largest environmental fraction (in DALYs, which combine years of life lost and years lived with disability for comparability of disease burden across diseases) include cardiovascular diseases, diarrhoeal diseases and lower respiratory infections. Ambient and household air pollution and water, sanitation and hygiene are the main environmental drivers of those diseases.

5. The evidence on quantitative links between health and environment has increased

A greater share of the estimates of the burden of disease attributable to the environment can now be determined using more robust methods than previously. Comparative risk assessment methods, which are based on the best exposure and/or exposure-response data, now cover 68% of deaths attributable to the environment, and 56% of DALYs (Figure ES12). In other words, about 8.6 million out of the total 12.6 million environment-attributable deaths are estimated on the basis of CRA methods and the remaining 4.1 million of epidemiological estimates using more limited data, transmission pathways and experts opinion. It should be noted that many of the potential health implications of climate change, acting through food supply and migration could unfortunately not be accounted for by the methods used in the report.

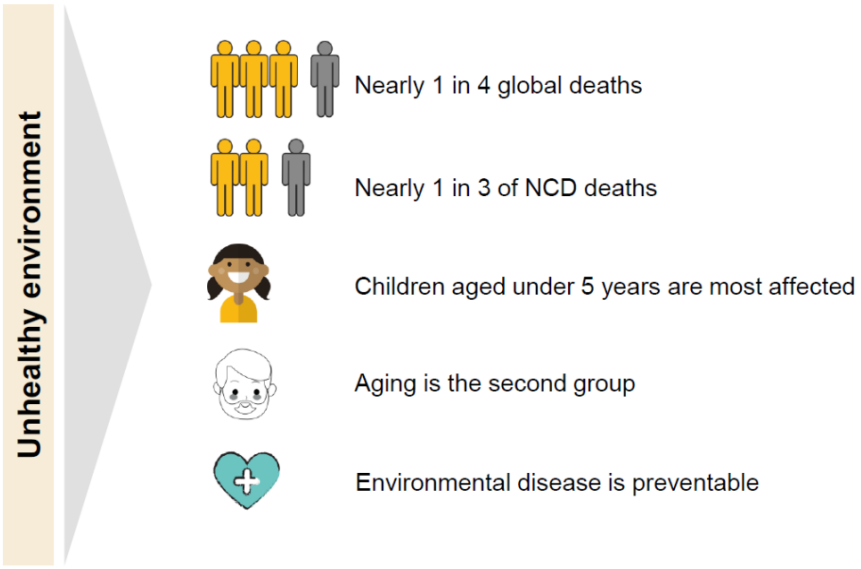
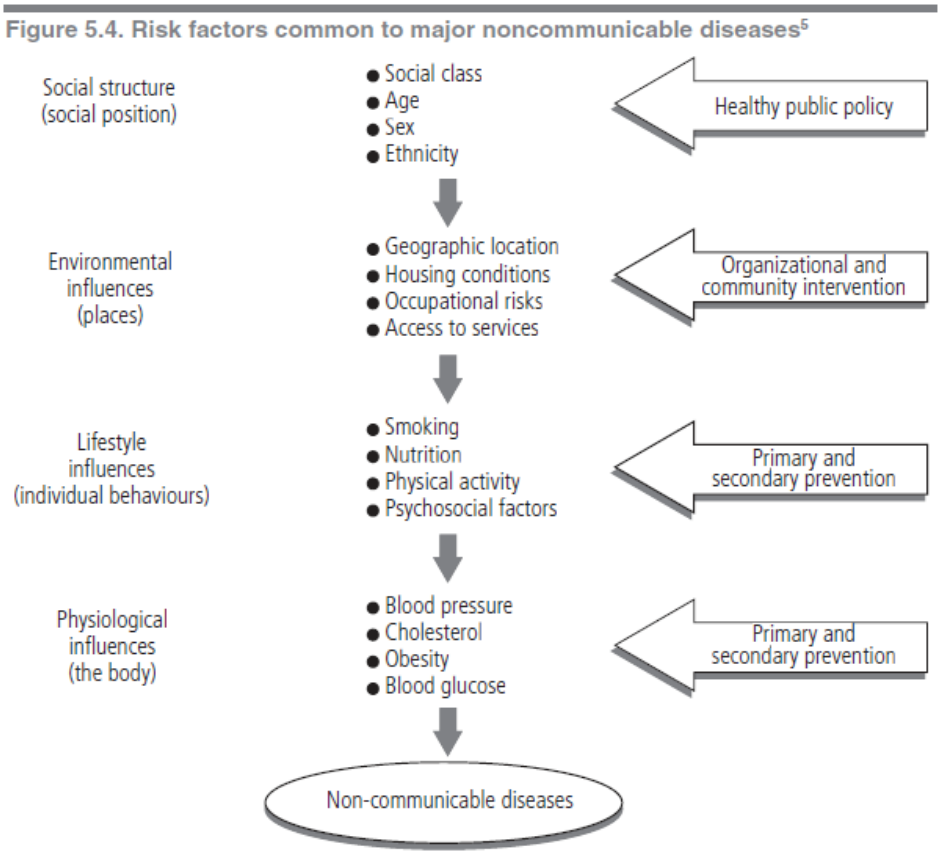

Unhealthy environment
Nearly 1 in 4 global deaths
Nearly 1 in 3 of NCD deaths
Children aged under 5 years are most affected
Aging is the second group
Environmental disease is preventable

Figure xx.xx A linkage between unhealthy environment and human health

(Source WHO, by Nutta Taneepanichskul, PhD)

Definition of Environmental Epidemiology

Environmental epidemiology is the study of the effect of physical, biological, and chemical factors on human health from the external environment through the examination of specific populations or communities exposed to different ambient environments. Environmental epidemiology seeks to clarify the relationship between physical, biological, and chemical factors and human health."



5300501 – Lecture # Session 15 Field Study

Before the field trip day, students are required to read 2 following articles to understand the basic concept of primary health care and Universal coverage scheme in Thailand

1. Rohde, J., Cousens, S., Chopra, M., Tangcharoensathien, V., Black, R., Bhutta, Z. A., & Lawn, J. E. (2008). 30 years after Alma-Ata: has primary health care worked in countries?. *The Lancet*, *372*(9642), 950-961.
2. Tangcharoensathien, V., Swasdiworn, W., Jongudomsuk, P., Srithamrongswat, S., Patcharanarumol, W., & Thammathat-aree, T. (2010). Universal coverage scheme in Thailand: equity outcomes and future agendas to meet challenges. *World health report*.

PowerPoint Handout